Unravelling the Evolution of Allosteric Regulation in 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate Synthase

---

A thesis submitted in partial fulfilment of
the requirements for the degree
of
Doctor of Philosophy in Biochemistry
in the Department of Chemistry
by Penelope Jane Cross

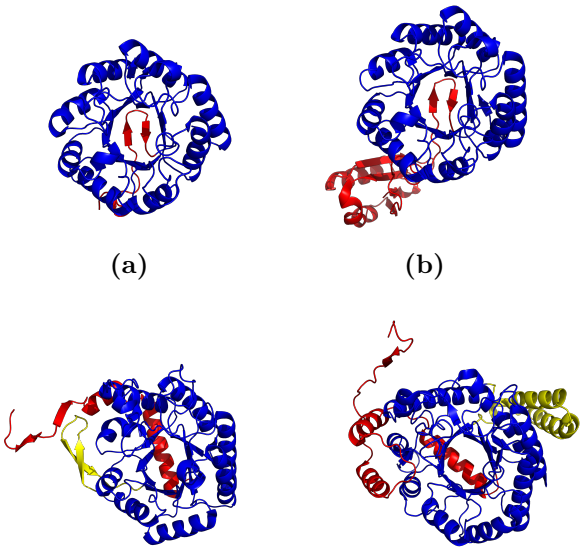
February 2012

Abstract

The enzyme 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (DAH7PS) catalyses the first reaction in the shikimate pathway, leading to the biosynthesis of aromatic compounds including the aromatic amino acids. The catalytic activity of DAH7PS is regulated through feedback inhibition and is the major control point for the pathway. DAH7PSs are divided into two families, type I and type II, based on molecular weight and amino acid sequence. Type I DAH7PSs can be further divided based on sequence similarity. All DAH7PS enzymes with their crystal structures solved share a basic $(\beta/\alpha)_8$-barrel fold in which the key catalytic components are housed. Furthermore, all structurally characterised DAH7PSs, except *Pyrococcus furiosus* DAH7PS (*Pfu*DAH7PS) and *Aeropyrum pernix* DAH7PS, have recruited extra structural motifs that are implicated in allosteric regulation. However, there are significant differences in the additional structural elements.

This thesis investigates the hypothesis that the diverse regulation strategies for controlling DAH7PS activity have evolved by domain recruitment, whereby regulatory domains have been added to the catalytic barrel.

Chapter 2 describes the functional characterisation of the type Iβ *Thermotoga maritima* DAH7PS (*Tma*DAH7PS), and the exploration of its response to inhibitors. The catalytic activity of *Tma*DAH7PS was found to be substantially inhibited by tyrosine (Tyr) and to a lesser extent, phenylalanine (Phe). The putative regulatory domain previously identified as a ferredoxin-like domain was recognised as an aspartate kinase-chorismate-mutase-tyrA



Variation of the structural elements decorating the basic barrel structure of the DAH7PS enzymes. (a) *Pyrococcus furiosis* DAH7PS, (b) *Thermotoga maritima* DAH7PS, (c) *Escherichia coli* DAH7PS and (d) *Mycobacterium tuberculosis* DAH7PS.

(prephenate dehydrogenase) or ACT domain.

Chapter 3 describes the characterisation of *Tma*DAH7PS with the N-terminal domain removed. The truncated enzyme was found to be more catalytically active than wild-type *Tma*DAH7PS and insensitive to inhibition by the aromatic amino acids, Tyr, Phe and tryptophan. Apart from the truncation of the ACT domain, the crystal structure of $^{\text{trunc}}$*Tma*DAH7PS showed no major changes to the monomer structure when compared to wild-type *Tma*DAH7PS. However, $^{\text{trunc}}$*Tma*DAH7PS crystallises as a dimer, unlike wild-type *Tma*DAH7PS.

In Chapter 4, the solution of the crystal structure of *Tma*DAH7PS with Tyr bound is presented. Tyr binding was shown to induce a significant conformational change, and Tyr is observed to bind at the interface between the ACT domains from two diagonally located monomers of the tetramer.

The major reorganisation of the regulatory domain with respect to the barrel observed in the crystal structure, was confirmed by small angle X-ray scattering. The closed conformation adopted by the protein on Tyr binding physically gates the neighbouring barrel and blocks substrate entry into the active site.

Chapter 5 explores the interactions between *Tma*DAH7PS and the allosteric inhibitor, Tyr. The residues His29 and Ser31, which form hydrogen bonds with the hydroxyl moiety of the Tyr ligand, were examined for their impact on the sensitivity and selectivity of the enzyme for the inhibitors Tyr and Phe. The hydroxyl side chain of Ser31 was found to be important for both the preferential inhibition by Tyr over Phe and the inhibitory mechanism. His29 (the hydrogen-bonding partner of Ser31) appears to play a secondary role in determining ligand selectivity and the relative positioning of these two residues is crucial to the inhibition of the enzyme.

Chapter 6 evaluates the transferability of allosteric control of catalytic activity. The ACT domain of *Tma*DAH7PS was fused onto the barrel of the unregulated *Pfu*DAH7PS. This chimeric enzyme was found to be catalytically active, inhibited by Tyr (although less sensitive) and preliminary crystallographic results show inhibition occurs via the same conformational change observed for wild-type *Tma*DAH7PS.

Acknowledgements

This thesis owes its existence to the help, support and inspiration of a number of people. Firstly, my crazy supervisor, Emily Parker whose sheer dedication and passion for science is infectious. You are someone to aspire to on both a personal and professional level but I will not miss following you to the bathroom to 'grab' some extra time or trying to do two days straight at the Tron with you. You are an incredible lady with tremendous stamina. Thank you for your guidance and friendship.

Thank you to Dr Renwick Dobson, Dr Sean Devenish and Dr Grant Pearce for letting me ask a million questions and always having both the time and patience to answer them. Also, thank you to Dr Andrew Muscroft-Taylor for your in-depth answers and spending your personal time to help me whenever you could.

Dr Jan Wikaira and Assoc. Prof. Richard Hartshorn, without your encouragement I would never have dreamed I was clever enough to do an Honours degree, let alone a PhD. Thank you for believing in me.

Thanks must go to the various agencies that have provided financial support over the years; to the University of Canterbury for a Doctoral Scholarship, the New Zealand Institute of Chemistry, the New Zealand Society for Biochemistry and Molecular Biology, and the New Zealand Federation of Graduate Women for funds to travel to various conferences. Thanks to the Biomolecular Interaction Centre for provision of well used (and loved) equipment and many sources of expert advice.

I owe tremendous thanks to my work husband, Timothy Allison (aka Sheldon), for support on both a professional and personal level. I can only hope I have had as much influence on your life as you have had on mine. A special thanks to Sebastian Reichau for sharing your desk with me during the earthquakes and many thanks to the other members of the Parker research group, especially Dr Aidan Harrison, for many 'interesting' conversations.

Thank you to my incredibly supportive friends, Marie, Megan, Skunk, Nana and Emma, some of whom have no idea what I have actually been doing but encourage and support me nonetheless.

Finally, thanks to my husband Corey and daughter Justice, who have been incredibly patient and supportive while watching a three year Bachelor degree turn into a nine year PhD. Your support has been invaluable and I truly could not have done this without you.

Contents

Abbreviations

A5P arabinose 5-phosphate

AEC anion-exchange chromatography

Ala alanine

Amp ampicillin

***Ape*DAH7PS** *Aeropyrum pernix* DAH7PS

AR analytical reagent

Arg arginine

bp base pair

BSA bovine serum albumin

***Bsu*DAH7PS** *Bacillus subtilis* DAH7PS

BTP 1,3-(tris(hydroxymethyl)-methylamino)propane

Cam chloramphenicol

CD circular dichroism

CM chorismate mutase

DAH7P 3-deoxy-D-*arabino*-heptulosonate 7-phosphate

DAH7PS 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase

DNA deoxyribonucleic acid

dNTP deoxyribonucleotide triphosphate

DSC differential scanning calorimetry

DTT dithiothreitol

E4P erythrose 4-phosphate

***Eco*DAH7PS** *E. coli* DAH7PS

EDTA ethylenediaminetetraacetic acid

EPSPS 5-enolpyruvyl shikimate 3-phosphate synthase

FL ferredoxin-like

FPLC fast protein liquid chromatography

G3P glycerol 3-phosphate

Gly glycine

GST glutathione S-transferase

HIC hydrophobic-interaction chromatography

His histidine

IMAC immobilised metal affinity chromatography

IPTG isopropyl-1-thiol-β-D-galactopyranoside

Kan kanamycin

KDO 3-deoxy-D-*manno*-2-octulosonic acid

KDO8P 3-deoxy-D-*manno*-octulosonate 8-phosphate

KDO8PS 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase

LB lysogeny broth

Lys lysine

MBP maltose binding protein

***Mtu*DAH7PS** *Mycobacterium tuberculosis* DAH7PS

MWCO molecular weight cut-off

ORF open reading frame

PAGE polyacrylamide gel electrophoresis

PCR polymerase chain reaction

PDB Protein Data Bank

PEP phosphoenolpyruvate

***Pfu*DAH7PS** *Pyrococcus furiosus* DAH7PS

***Pgi*DAH7PS** *Porphyromonas gingivalis* DAH7PS

Phe phenylalanine

pI isoelectric point

RMSD root-mean-square deviation

SAXS small angle X-ray scattering

***Sce*DAH7PS** *Saccharomyces cerevisiae* DAH7PS

SDS sodium dodecyl sulfate

SEC size-exclusion chromatography

Ser serine

SOC super optimal broth

Spec spectinomycin

TAE tris base-acetic acid-EDTA

TCEP tris(2-carboxyethyl)phosphine

TE tris-EDTA

TEV Tobacco Etch Virus

TIM triosephosphate isomerase

***Tma*DAH7PS** *Thermotoga maritima* DAH7PS

Trp tryptophan

Tyr tyrosine

UV ultra-violet

List of figures

List of tables

Publications

A subset of this thesis has been published in the following publication:

Cross, Penelope J.; Dobson, Renwick C.J.; Patchett, Mark L.; Parker, Emily J. Tyrosine latching of a regulatory gate affords allosteric control of aromatic amino acid biosynthesis. *Journal of Biological Chemistry* **2011**, (286), 10216–10224.

Chapter 1

Introduction

1.1 The Shikimate Pathway

Organisms differ in their capability to carry out the chemical reactions required to produce the amino acids, the constituents of proteins. Most microorganisms are proficient in all such biosyntheses, whereas animals are able to manufacture approximately only half of the amino acids. Among the amino acids that cannot be biosynthesised by animals are the three aromatic amino acids, tyrosine (Tyr), phenylalanine (Phe) and tryptophan (Trp). In microorganisms and plants these important compounds are produced via the shikimate pathway.[1,2]

The shikimate pathway is a series of seven enzyme-catalysed reactions (Figure 1.1) responsible for the biosynthesis of chorismate, the precursor for many important aromatic compounds. Branching of the shikimate pathway occurs after the formation of chorismate (derived from the Greek term meaning "fork"), where one branch leads to anthranilate and thereon to Trp, and the other branch leads to prephenate, which is a precursor for both Tyr and Phe. Further branch points lead to the formation of other aromatic and pre-aromatic compounds including various quinones, folic acid and vitamins E and K.[3]

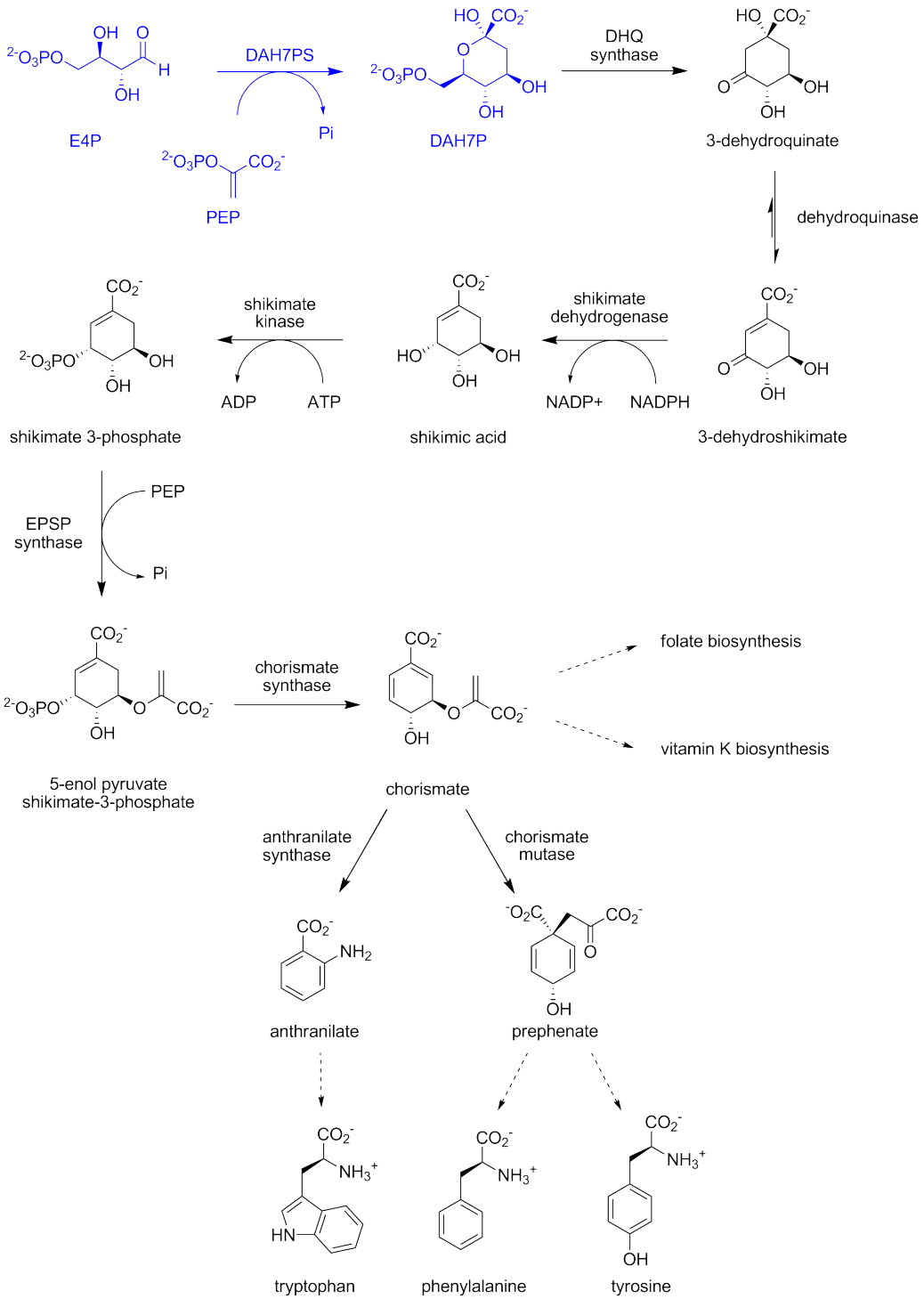


Figure 1.1: Overview of the seven enzyme-catalysed reactions of the shikimate pathway leading to chorismate before the pathway branches to form important aromatic compounds. The reaction catalysed by DAH7PS is coloured blue.

The shikimate pathway operates in bacteria, plants and fungi and has more recently been shown to operate in apicomplexan parasites.[4] As the pathway is absent in mammals, considerable interest has been expressed in the pathway as a possible target for antimicrobial agents and herbicides.[5] Successful inhibition of the shikimate pathway has been demonstrated by the broad-spectrum herbicide glyphosate, which inhibits the sixth enzyme of the pathway, 5-enolpyruvyl shikimate 3-phosphate synthase (EPSPS).[6,7]

1.1.1 Discovery of the Shikimate Pathway

Shikimic acid was the first intermediate of the shikimate pathway (Figure 1.1) to be discovered and was first isolated from *Illicium religiosum*,[8] the Japanese name shikimi-no-ki (anise tree) of which the shikimate pathway draws its name.[9] It was not until the 1950s when Davis[1] obtained a series of *E. coli* mutants that could only grow by supplementation with aromatic substances, that shikimate, due to its accumulation, was implicated as an intermediate in the biosynthesis of the aromatic amino acids and other essential substances.[10,11] The other intermediates of the pathway were elucidated using ultra-violet (UV)-light induced mutants of *E. coli*, *Aerobacter aerogenes* and *Neurospora*.[2,12–16] Isotope tracer experiments using ^{14}C-D-glucose showed that shikimic acid is derived from phosphoenolpyruvate (PEP), a three-carbon product from glycolysis, and the remaining four carbons from erythrose 4-phosphate (E4P) provided by the pentose phosphate pathway.[17]

1.2 Introduction to 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate Synthase

The first enzyme of the shikimate pathway is 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (DAH7PS) (EC 2.5.1.54), which catalyses an

aldol-like condensation reaction between the pentose phosphate pathway intermediate, E4P and the glycolytic intermediate, PEP.[18] This generates 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAH7P), a seven-carbon sugar, with the release of inorganic phosphate (Pi) (Figures 1.1 and 1.2).

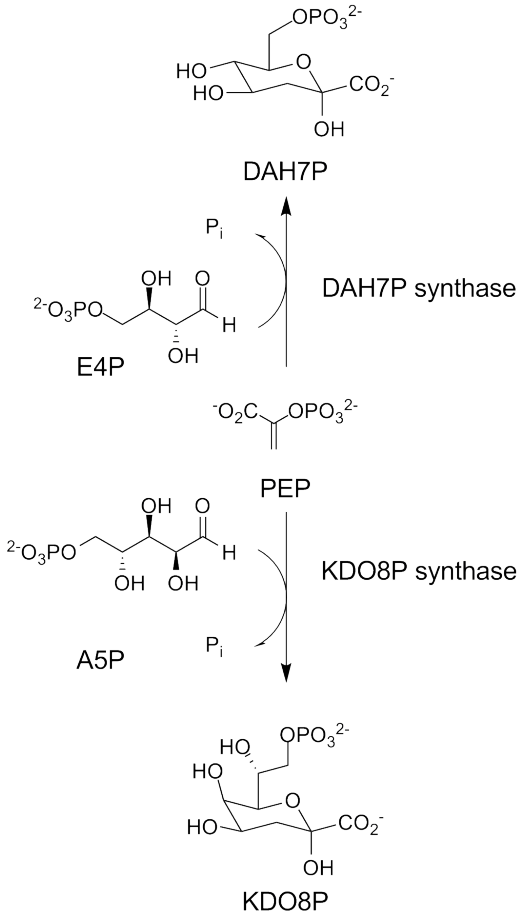


Figure 1.2: Condensation of PEP with phosphorylated monosaccharides.

Regulation of the catalytic activity of DAH7PS has been shown to be an important mechanism for control of cellular levels of aromatic compounds in microorganisms and plants.[19] The overall flux through the shikimate pathway is controlled via inhibition of DAH7PS, usually mediated by the end products of the pathway, the aromatic amino acids.[20]

Other enzymes found at branch points throughout the shikimate pathway are also inhibited by the pathway end products. In *B. subtilis*, prephenate

dehydratase, prephenate dehydrogenase and anthranilate synthetase, rather than DAH7PS, are reported to be inhibited by Phe, Tyr and Trp respectively.[21,22] *B. subtilis* DAH7PS (*Bsu*DAH7PS) itself, is feedback regulated by chorismate.[23]

1.2.1 Classification of DAH7PSs

DAH7PSs from different species can be divided into two distinct homologous families based on their molecular weight and amino acid sequence. These families have been referred to as type I and II according to the notation of Jensen.[24] The type I family of DAH7PSs can be further divided into type Iα and Iβ subfamilies based solely on sequence similarity.

The first and most extensively studied group of DAH7PS enzymes belong to the type Iα subfamily, and this group includes the well-characterised enzymes from *E. coli*[25–32] and *S. cerevisiae*.[33,34] With molecular masses below 40 kDa, the type I DAH7PSs are smaller than their type II orthologues.

The levels of overall sequence identity between individual members of the type Iα and type Iβ subfamilies are relatively low (approximately 18 percent), suggesting little homology between the two subfamilies.[35] Enzymes belonging to the type Iβ DAH7PS subfamily are generally smaller than their type Iα counterparts (30 kDa to 38 kDa in comparison to approximately 40 kDa) and are more closely related in sequence to the 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase (KDO8PS) enzymes (denoted class $I\beta_K$) than the type Iα subfamily (Figure 1.3).[24]

KDO8PS catalyses a reaction analogous to that catalysed by DAH7PS, condensing PEP with a five-carbon sugar, arabinose 5-phosphate (A5P), to form 3-deoxy-D-*manno*-octulosonate 8-phosphate (KDO8P) and inorganic phosphate (Figure 1.2). The reaction catalysed by KDO8PS is the first committed step in the formation of 3-deoxy-D-*manno*-2-octulosonic acid (KDO), an essential component of the lipopolysaccharide layer in the cell walls of Gram-negative bacteria. Formation of the lipopolysaccharide layer is import-

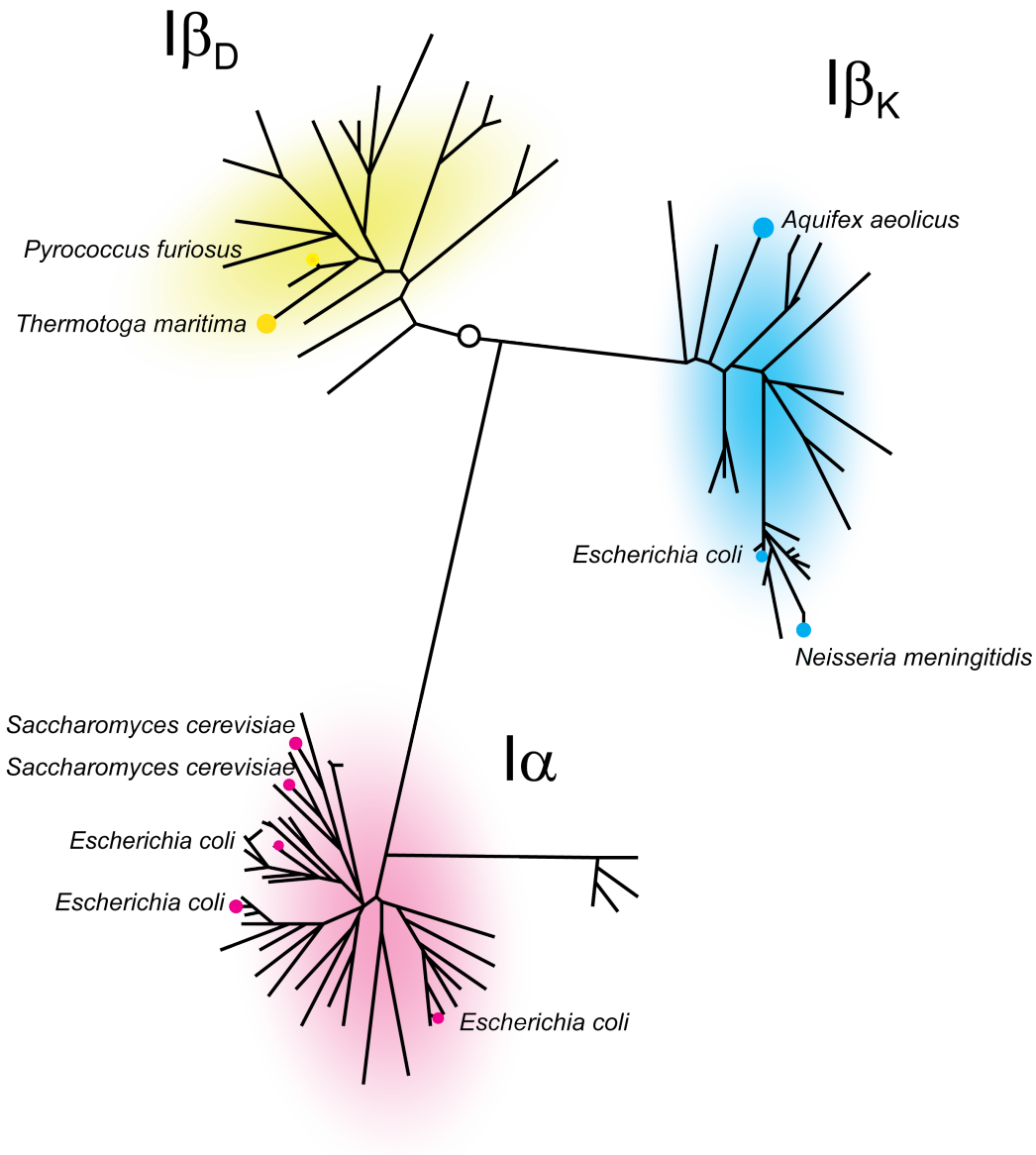


Figure 1.3: Phylogenetic homology analysis, modified from Jensen[24] comparing Iα and Iβ enzymes. The Iα DAH7PS enzymes are feedback regulated. The Iβ subfamily encompasses the type Iβ DAH7PS (Iβ$_D$) (both regulated (*Tma*DAH7PS) and unregulated (*Pfu*DAH7PS)), and the closely related KDO8PS enzymes (Iβ$_K$). The hypothetical ancestral root is denoted by a white circle.

ant for both growth and virulence of Gram-negative bacteria, identifying KDO8PS as a possible target for anti-bacterial drug design.[36,37]

Despite the lack of overall sequence similarity between the Iα, $I\beta_D$ and $I\beta_K$ enzymes, multiple sequence alignments indicate that they comprise a family of homologues that share a number of conserved residues,[35] suggesting common evolutionary ancestry. This relationship has been further cemented by the determination of several crystal structures of type Iα and Iβ DAH7PSs, and KDO8PSs, which show that all three groups of enzymes are structurally similar (refer to section 1.2.4).[25,34,38–40]

Type II DAH7PSs, with a molecular weight of approximately 54 kDa, are larger than their type I orthologues and were originally found in higher plants where they were thought to function in aromatic secondary metabolism.[9,41] A limited number of microbial enzymes were believed to exist within this family. However, as more microbial genes have been sequenced, evidence suggests that the type II DAH7PS family actually consists of a small set of plant enzymes within a diverse set of microbial enzymes.[42] A small number of organisms have been found to contain both type I and type II enzymes including *Xanthomonas campestris*, *Pseudomonas aeruginosa*, *Stigmatella aurantiaca*, *Amycolatopsis methanolica* and *Amycolatopsis mediterranei*[43,44] and several of the type II DAH7PSs from these organisms have been shown to be required for the biosynthesis of specific secondary metabolites.[44,45] Evidence that type II DAH7PSs function in aromatic amino acid biosynthesis and therefore primary metabolism, results from the presence of only type II enzymes in a number of species including *Helicobacter pylori*, *Streptomyces*, *Corynebacterium diphtheriae*, *Novosphingobium aromaticivorans* and several *Mycobacteria*.

Microbial type II enzymes share less than ten percent amino acid sequence identity with the type I enzymes. However, structural information (as discussed in section 1.2.5) reveals similar order and spacing of active-site residues between the type I and type II enzymes.[46,47]

1.2.2 Catalytic Mechanism

The reaction mechanism of DAH7PS has been subject to extensive structural and kinetic studies and many of the key details have been elucidated (Figure 1.4).[26,27,34,48–50] Analysis of the product formed when using deuterated PEP as a substrate shows the reaction proceeds in a stereospecific manner, where the *si* face of the alkene functionality of PEP attacks the *re* face of the aldehyde group of E4P.[51] Unusually, rather than utilising the high energy contained in the phosphate ester bond (ΔG° of $-62\,\mathrm{kJ\ mol^{-1}}$) by cleavage of the P-O bond of PEP,[52] DAH7PS cleaves the C-O bond of PEP.[5] This requires that water must act as a nucleophile at C2 of PEP at some stage of the reaction.

Early studies using *E. coli* DAH7PS (*Eco*DAH7PS) indicated that a ping-pong mechanism may operate in DAH7PS during catalysis.[53] Later kinetic studies using lower concentrations of substrates and more highly

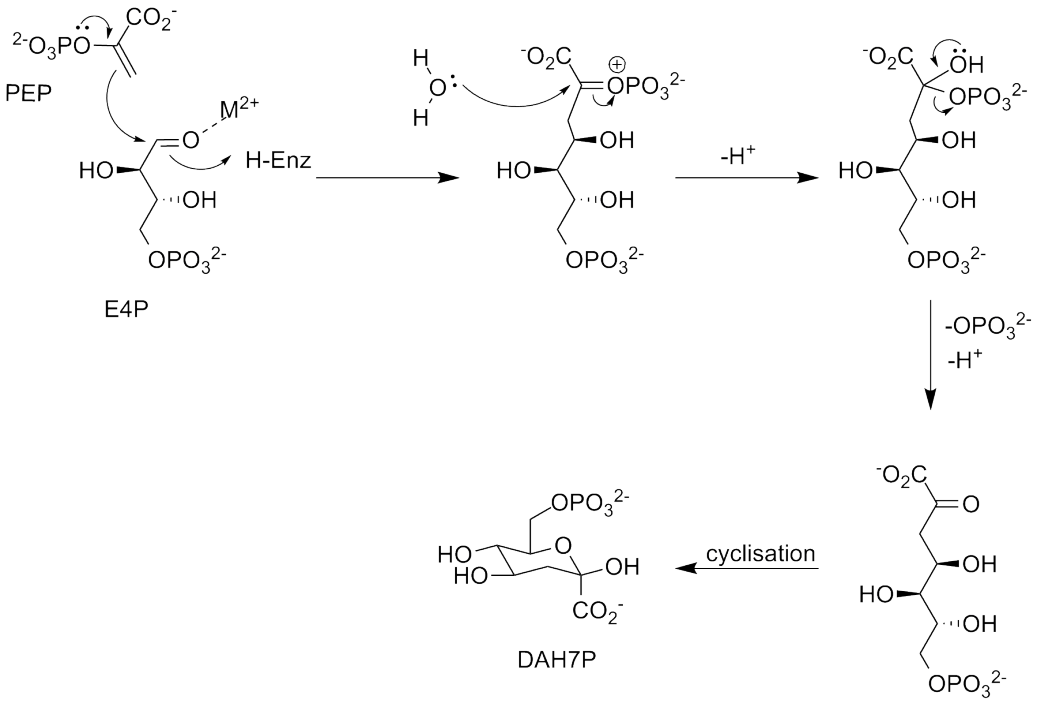


Figure 1.4: Proposed mechanism of condensation of PEP with E4P catalysed by DAH7PS.

purified enzyme are consistent with the reaction proceeding via an ordered sequential mechanism, where PEP binds to the enzyme first, followed by E4P, phosphate release, and finally DAH7P release.[50]

The first step of the reaction is thought to be initiated by the C3 of PEP attacking C1 of E4P, which is activated by coordination of the carbonyl group of E4P to an essential divalent metal ion. This addition step of the reaction occurs in a stereospecific manner and the resulting oxocarbenium ion is then attacked by water, producing a bisphosphate hemiketal intermediate. This tetrahedral intermediate then undergoes C-O bond cleavage, releasing phosphate and forming the C2 carbonyl of DAH7P. The product is thought to be released from the enzyme before formation of the more stable cyclic pyranose form.[54]

1.2.3 Metal Activation

Unlike the related KDO8PS enzymes, which can be metal dependent or metal independent, all DAH7PS enzymes that have been characterised to date are metal dependent. A variety of divalent metal ions have been proposed to reside in the active site under physiological conditions including iron,[55,56] cobalt[30,57] and copper.[58] Some DAH7PS enzymes are reported to be inactivated by the addition of a metal chelator, such as ethylenediaminetetraacetic acid (EDTA), with restoration of catalytic activity occurring with addition of a divalent metal ion.[30,58–61] On the contrary, other DAH7PS enzymes have been shown to be unaffected by EDTA or are unable to be reactivated by addition of excess metal ions.[28,29,62]

Generally, DAH7PS enzymes can be activated by a number of different divalent metal ions including Mn^{2+}, Cd^{2+}, Co^{2+}, Cu^{2+}, Zn^{2+} and Fe^{2+}. The variation within the metal-activation series does not appear to correlate with the subfamily of DAH7PS and although the exact identity of the metal ion remains undefined, it is likely to be dependent on bioavailability.

Two roles have been proposed for the metal ion. The most likely role is

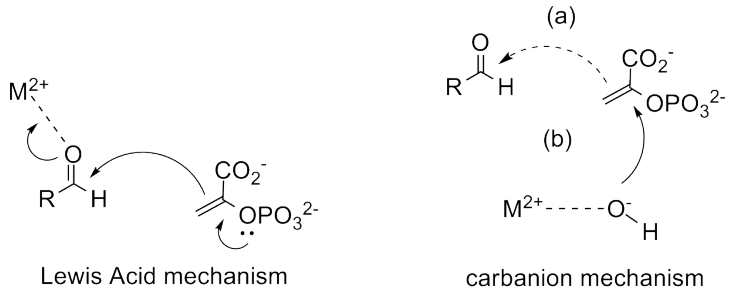


Figure 1.5: The two possible roles for the metal ion in DAH7PS.

as a Lewis acid, in which the metal functions to encourage attack of E4P by PEP by withdrawing electron density from the E4P carbonyl group (Figure 1.5). The alternative role is that the metal ion reduces the pKa of the nucleophilic water through coordination so as to provide a hydroxide ion for the attack at C2 of PEP. In this mechanism, C-O bond formation at C2 of PEP would precede C-C bond formation between PEP and E4P, and the reaction would proceed via a carbanion intermediate or transition state with carbanion character at C3 of PEP.

There are a number of problems associated with this alternative carbanion mechanism. Firstly, in order for hydroxide to attack, C2 of PEP requires unprecedented electrophilic character.[63] In addition, this mechanism requires the presence of a resonance unstabilised carbanion that would likely be quenched by hydrogen abstraction from another species within the immediate environment, such as water or the carboxylate or ammonium containing side chains of the protein itself.

1.2.4 Structure of DAH7PSs

Type Iα DAH7PSs

The most well characterised examples of DAH7PSs are those from *E. coli*. The genome of *E. coli* encodes three isozymes of DAH7PS, which are all members of the Iα subfamily. These enzymes are all approximately the same molecular

weight (approximately 38 kDa) and are similar in sequence.[25] However, each isozyme is specifically inhibited by one of the aromatic amino acids.

The Phe-sensitive *Eco*DAH7PS, *Eco*DAH7PS(Phe), was the first structure of a DAH7PS to be solved by X-ray crystallography.[25] This structure, and subsequent structural studies of other DAH7PSs,[34,38,40,46,47,64] show that the tertiary structure of all DAH7PS enzymes is comprised of a $(\beta/\alpha)_8$ or triosephosphate isomerase (TIM)-barrel monomer fold,[27] a common fold for globular proteins. As is typical for TIM-barrel enzymes, the active site of DAH7PS is located at the C-terminal end of the barrel.[65] The loops connecting helices and strands located at the N-terminal end of the barrel (α-β loops) are shorter than the loops at the C-terminal end of the barrel (β-α loops). These extended loops provide the residues that form the active site.

The *Eco*DAH7PS(Phe) crystal structure also reveals the enhancement of the core barrel catalytic domain with an N-terminal extension (strand β0 followed by two helices α00 and α0), and an internal β6a/β6b sheet inserted before the β6 strand of the barrel (Figure 1.6).[25] Both of these extensions are located outside the core barrel and are implicated in the allosteric inhibition of the enzyme.[32,33]

*Eco*DAH7PS(Phe) crystallises as a tetramer, consisting of a dimer of two tight dimers (Figure 1.7). Within the dimers, the formation of an intermolecular β-sheet involving the N-terminal β0 strand of one monomer with the βa-βb strands of a neighbouring subunit, enhances the association.[25]

S. cerevisiae expresses two isozymes of DAH7PS, which are also categorised into the type Iα subfamily of DAH7PS enzymes. As for the *Eco*DAH7PS(Phe) enzyme, the Tyr-repressible isozyme, *S. cerevisiae* DAH7PS (*Sce*DAH7PS)(Tyr), also crystallises as a tetramer. Despite originating from different organisms, *Eco*DAH7PS(Phe) and *Sce*DAH7PS(Tyr) are very similar enzymes in terms of sequence (two-hundred conserved residues) and of monomer folds (root-mean-square deviation (RMSD) of 0.6 Å between equivalent Cα atoms) (Figure 1.6).

Although one of the two dimers between *Eco*DAH7PS(Phe) and *Sce*-

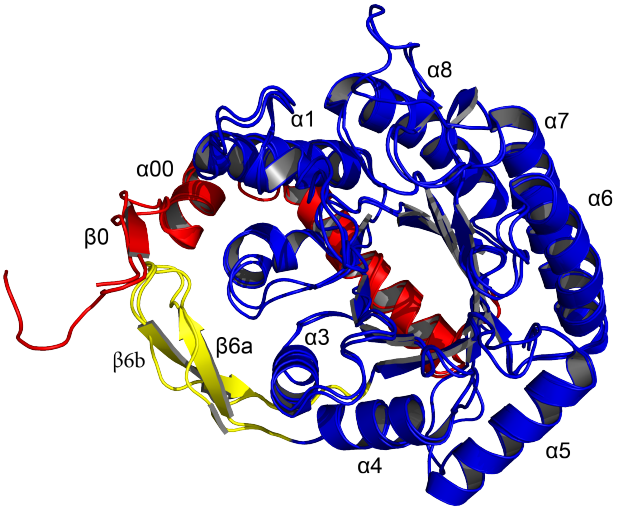
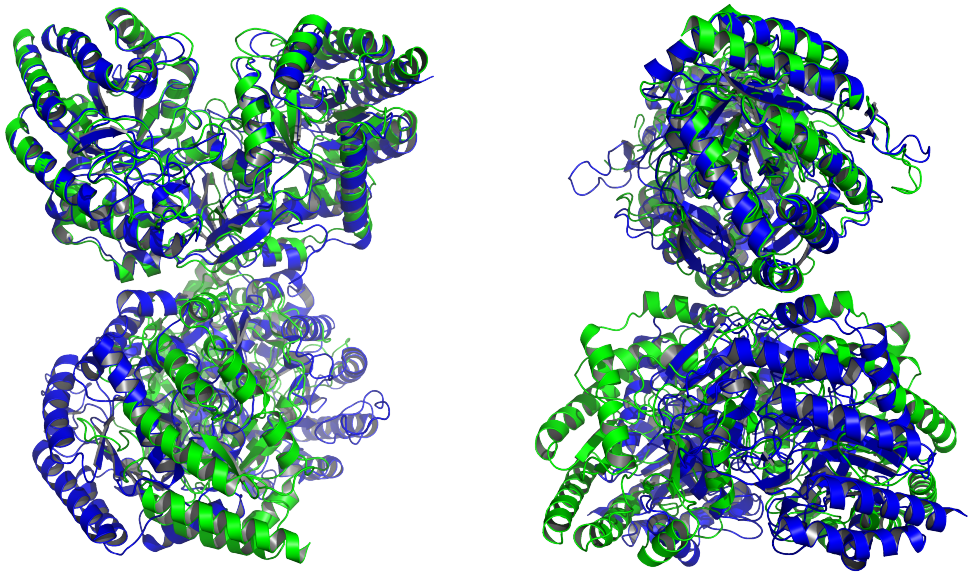


Figure 1.6: Comparison of the monomeric structure of *Eco*DAH7PS (PDB code 1KFL) and *Sce*DAH7PS (PDB code 1HFB). The core barrel is coloured blue and N-terminal and internal extensions are coloured red and yellow, respectively.



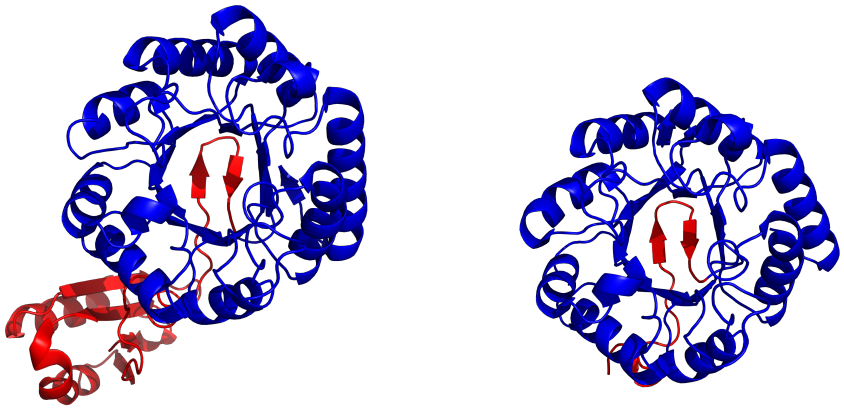
(a) **(b)**

Figure 1.7: Superposition of the tetrameric structure of *Eco*DAH7PS (PDB code 1KFL) coloured blue and *Sce*DAH7PS (PDB code 1HFB) coloured green.

DAH7PS(Tyr) is essentially superimposable (Figure 1.6), the interfaces that form the tetramers are slightly twisted relative to each other (Figure 1.7). Furthermore, in *Sce*DAH7PS(Tyr) the tetrameric association is purely hydrophobic whereas electrostatic and hydrogen-bonding interactions play a more important role in *Eco*DAH7PS(Phe). Regardless of these subtle differences, the type Iα subfamily of DAH7PS appear to be defined by the similarity of extensions and insertions to the core barrel.

Type Iβ DAH7PSs

The first Iβ DAH7PS crystal structure solved was of *T. maritima* DAH7PS (*Tma*DAH7PS) (Figure 1.8a).[39] Again, the structure incorporates a core $(\beta/\alpha)_8$ barrel, and besides the eight strands and helices that form the barrel, the monomer units include a two-stranded antiparallel β-hairpin that covers the N-terminal end of the barrel. The core barrel structure is connected at the N terminus via a linker to another compact domain. The extra domain, with a distinctive βαββαβ fold, is described as a ferredoxin-like (FL) domain. It was speculated that this FL domain is involved in the feedback regulation



(a) **(b)**

Figure 1.8: Monomeric structure of (a) *Tma*DAH7PS (PDB code 1RZM) and (b) *Pfu*DAH7PS (PDB code 1ZCO) with the core barrels coloured in blue and the N-terminal extensions in red.

of *Tma*DAH7PS.[39]

The crystal structure of *Tma*DAH7PS shows that two dimers associate to form a tetramer by extensive interactions between the $(\beta/\alpha)_8$ barrels. The segments of the barrel involved in tetramer formation are helices α5–α8 and loops β5-α5, β6-α6 and β7-α7. The FL domains and barrels do not interact other than through their direct linkages and do not participate in the contacts between subunits.

A common dimeric unit is shared between the type Iα and type Iβ subfamilies of the DAH7PSs. However, although both subfamilies crystallise as tetramers, the tetramers are formed by utilising completely different structural elements.[25,34,39,48] This means that although dimers can be superimposed, the

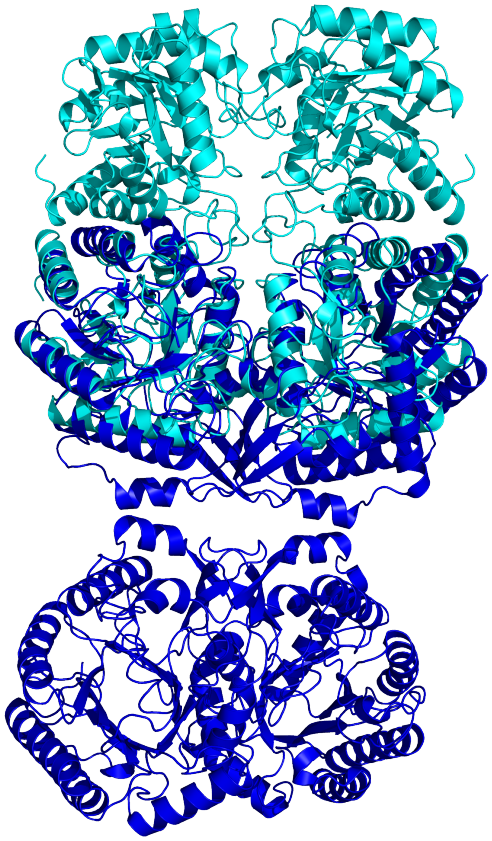


Figure 1.9: Superposition of the type Iα *Eco*DAH7PS (PDB code 1GG1), in blue, onto the type Iβ *Pfu*DAH7PS (PDB code 1ZCO), in cyan.

tetramers cannot (Figure 1.9).

Resembling *Tma*DAH7PS, *P. furiosus* DAH7PS (*Pfu*DAH7PS) (another type Iβ DAH7PS) consists of a $(\beta/\alpha)_8$ barrel with the N-terminal end of the barrel closed off by a two-stranded β-hairpin (Figure 1.8b).[48] This enzyme has no extra structural elements, meaning the monomer unit is composed of only the core catalytic barrel. Regardless of this difference, *Pfu*DAH7PS and *Tma*DAH7PS share the same tetramer structure in the crystalline form.

The crystal structure of *A. pernix* DAH7PS (*Ape*DAH7PS) has been solved and its molecular description was recently published.[40] *Ape*DAH7PS shares a similar unembellished barrel structure to *Pfu*DAH7PS and forms the same tetramer structure as both *Pfu*DAH7PS and *Tma*DAH7PS in the crystalline form.

Type II DAH7PSs

Of the type II family, only the *M. tuberculosis* DAH7PS (*Mtu*DAH7PS) crystal structure has been solved. Prior to this structure, it was unclear whether the type II enzymes would be structurally related to the other DAH7PSs. The *Mtu*DAH7PS monomer shares the same core-barrel fold as all other structurally characterised DAH7PS enzymes (Figure 1.10).[46] Two major additions decorate the barrel that are quite distinct from the additions to the type Iα and type Iβ enzymes. An N-terminal extension comprises of a β-strand followed by three helices (α0a, α0b, and α0c), the last of which closes off the N-terminal end of the barrel. The other addition is formed by two helices (α2a and α2b), which extend the α2-β3 connecting loop and pack against the outside of the barrel. Again, these extensions are implicated in allostery of the enzyme.[66]

Although comparison of the core barrel units of *Mtu*DAH7PS and the type Iα DAH7PSs suggest the same spatial arrangement of active-site residues, the quaternary structure of *Mtu*DAH7PS is distinct. The tetramer structure of *Mtu*DAH7PS is composed of two tightly associated dimers, however neither

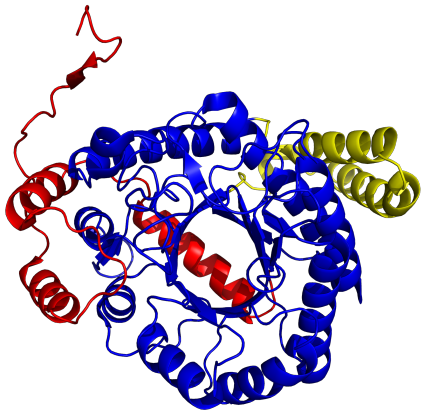


Figure 1.10: Monomeric structure of the type II *Mtu*DAH7PS (PDB code 3NUD). The core barrel is coloured blue, N-terminal extension red and the internal extension yellow.

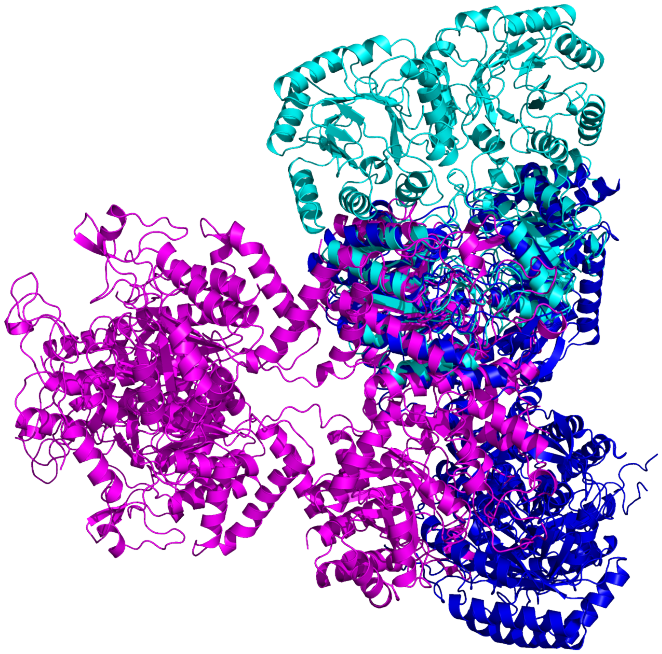


Figure 1.11: Comparison of the quaternary structures of DAH7PS enzymes. Superposition of the type II *Mtu*DAH7PS tetramer (pink), onto the type Iα *Eco*DAH7PS (cyan) and type Iβ *Pfu*DAH7PS (blue).

the dimer nor the monomer-monomer interface shares any commonality with any subunit interface found in other DAH7PS structures (Figure 1.11).

1.2.5 Architecture of the Active Site

Despite the low overall sequence similarity between the type Iα, Iβ and type II DAH7PS enzymes, both the identity and positioning of several key residues involved in the catalytic activity of these enzymes are absolutely conserved (Figure 1.12).

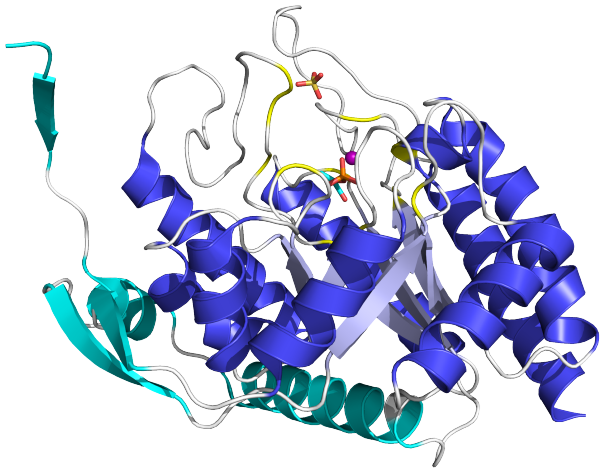


Figure 1.12: Side-on view of the architecture of the DAH7PS active site as illustrated by *Eco*DAH7PS(Phe) (PDB code 1N8F). The β-strands and α-helices belonging to the barrel are coloured light blue and blue, respectively. Non-barrel regions (cyan), PEP (carbon atoms are coloured cyan), manganese (magenta sphere) and a sulfate ion (yellow sticks) are indicated. The sulfate ion can be taken as an estimated position of the phosphate group of E4P. Protein residues established for ligand and metal binding are coloured yellow.

Metal Binding Site

All DAH7PSs require a divalent metal ion for catalytic activity. The metal co-ordination site is located near the C-terminal end of the barrel[25] and residues contributing to the co-ordination of the metal ion are absolutely

conserved across all DAH7PS families. In the structure of *Eco*DAH7PS, the metal ion, Mn^{2+}, is held in a distorted bi-pyramidal geometry, coordinated to the enzyme via residues positioned on the short loops β1-α1 (Cys61), β7-α7 (His268), strand β8 (Glu302) and on the long loop β8-α8 (Asp326) (Figure 1.13).

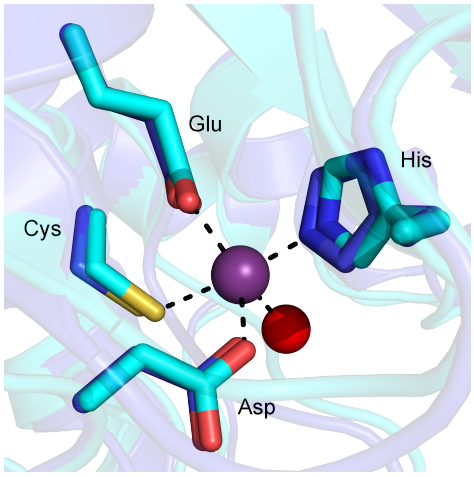


Figure 1.13: Overlay of the metal binding sites of *Eco*DAH7PS (PDB code 1N8F), coloured cyan and *Sce*DAH7PS (PDB code 1OF6), coloured blue. The metal ion and water molecule are shown as a purple and red sphere respectively.

PEP Binding Site

The PEP binding site provides a series of non-covalent interactions between the enzyme and PEP phosphate and carboxylate groups (Figure 1.14).[27,34] PEP is positioned in the active site with hydrogen-bonding and salt-bridge interactions mediated by three arginine (Arg) and two lysine (Lys) residues (*Eco*DAH7PS numbering Arg165, Lys196, Lys97, Arg92 and Arg234), all of which are conserved across all DAH7PS families.

E4P Binding Site

An absolutely conserved KPR(T/S) (*Eco*DAH7PS numbering Lys97, Pro98, Arg99, Thr100) motif is found across all DAH7PS enzymes and forms a

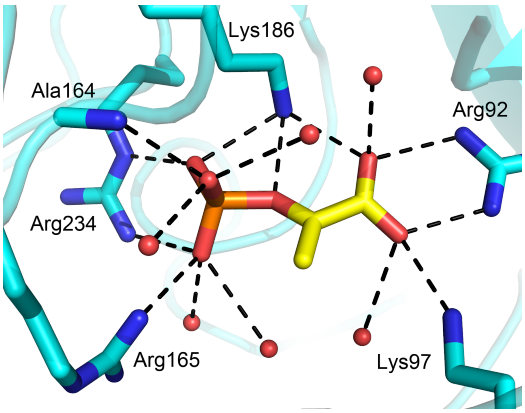


Figure 1.14: PEP binding site of *Eco*DAH7PS(Phe) (PDB code 1N8F). Dashed lines indicate hydrogen-bonding and salt-bridge interactions between PEP and the surrounding water molecules and protein residues.

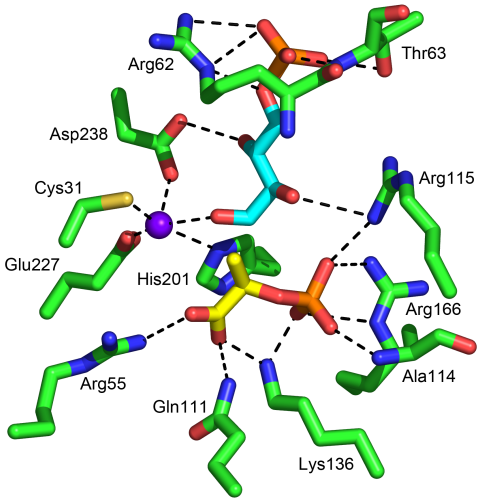


Figure 1.15: The binding site of E4P as modelled into the active site of *Pfu*DAH7PS.[48] E4P is shown as yellow sticks, PEP as cyan sticks and the metal ion, Mn^{2+}, as a purple sphere.

binding pocket for the enzyme's second substrate, E4P.

The crystal structure of *Tma*DAH7PS was solved in complex with all required reaction components, Cd^{2+}, PEP and E4P.[39] The presence of the two substrates was observed due to the slow catalysis by the hyperthermophilic enzyme at low temperatures. However, E4P was located at a site distant to PEP where the reaction is not feasible. While no plausible structures of enzyme-PEP-E4P complexes exist, the structure of *Sce*DAH7PS has been determined to high resolution with metal ion (Co^{2+}), PEP and glycerol 3-phosphate (G3P) bound.[34] Information from this structure has been used to infer the binding site of E4P, and also to model E4P into the active site of *Pfu*DAH7PS (Figure 1.15).[48]

Some DAH7PS enzymes are tolerant to the substitution of E4P with alternative aldehydes, with deoxy and five-carbon phosphorylated substrates being accepted as alternatives. The type Iβ enzymes show the greatest substrate promiscuity.[23,47,48,64,67–70]

1.3 Enzyme Regulation

1.3.1 Introduction

Acting in sequence, enzymes form metabolic pathways by which nutrient molecules are degraded, energy is released and is converted into useful, usable forms, and precursors are used and transformed into the thousands of distinctive biomolecules found to exist in a living cell at any one time.[71–73] Situated at key junctions are regulatory enzymes, which are able to sense the metabolic needs of the cell and adjust their catalytic activity accordingly.[74] The response of these enzymes ensures the integration of the diverse and divergent metabolic activities of the cell.

Metabolic pathways can be controlled at three levels, namely by the amount of enzymes available, and through modulation of enzyme activity and

substrate availability. Therefore, there are a number of different mechanisms governing the rates of enzyme catalysis including the control of transcription of deoxyribonucleic acid (DNA) and translation into protein, compartmentalisation,[75] post-translational modifications,[76] localisation[77] and allostery.[78]

1.3.2 Transcriptional Regulation

Although the major mechanism for regulation of DAH7PS activity is via feedback inhibition and the direct control of this enzyme's catalytic activity, other regulatory methods are also utilised. The three isozymes of *Eco*-DAH7PS are also regulated through repression of transcriptional activity. The expression of genes encoding for *Eco*DAH7PS(Phe) (*aro*G) and *Eco*-DAH7PS(Tyr) (*aro*F), appear to be controlled by the Tyr repressor, *tyr*R gene product,[79,80] when in combination with Phe and Tyr respectively.[81]

The regulation of the expression of *aro*H, the structural gene for the Trp-sensitive isozyme of *Eco*DAH7PS, *Eco*DAH7PS(Trp), is subject to control by the Trp repressor, *trp*R and by Trp itself.[82,83] When intracellular levels of Trp are low, the Trp repressor exists predominantly in a dissociated or aporepressor form (free of bound corepressor), which has a low affinity for the operator DNA.[84] However, when the level of Trp is elevated, due to increased biosynthesis or transport, the aporepressor form of the Trp repressor binds to its corepressor, Trp, to form an active repressor complex. This complex then binds the respective operator DNAs to repress Trp, *trp*R and *aro*H operon expression.[82] *In vitro* DNA binding experiments have shown that the Trp repressor binds at a regulatory site upstream of the *aro*H gene, preventing transcription of the *aro*H gene by interfering with polymerase binding.[82,83]

1.3.3 Allostery

The allosteric regulation of DAH7PS is one of the major mechanisms for control of the shikimate pathway.[20] In a general sense, allostery can be defined as the phenomenon where the activity or function of one part of a protein

is effected and affected by a binding event at a remote or allosteric site (Figure 1.16). Controlling enzyme activity permits the fine tuning necessary to balance the rate of synthesis of end products of the pathway with the production of subsequent larger macromolecules, and prevents the organism from expending unnecessary resources, such as nitrogen and energy, on end products that are not required.

In 1965, Monod, Wyman and Changeux summarised the properties of a number of allosteric enzymes resulting in their plausible model on the nature of allosteric transition (known as the MWC model, Figure 1.17).[85] Since then, the description of numerous enzymes had led to the concept that allostery is a common theme in enzyme regulation.

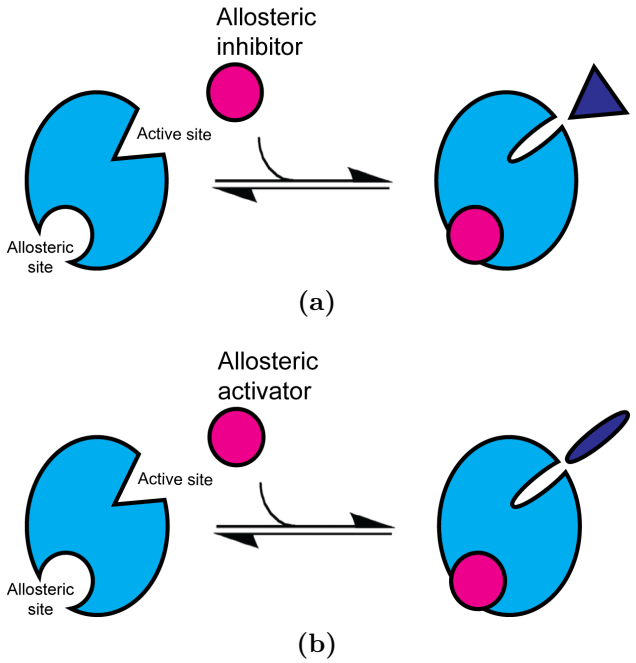


Figure 1.16: Cartoon representation of a monomeric, allosterically-regulated enzyme. The binding of an allosteric effector alters the active site or substrate binding site in either an (a) unfavourable or (b) favourable way, causing either a decreased affinity for substrate binding and therefore decreased catalytic activity or increase in catalytic efficiency resulting from increased substrate affinity, respectively.

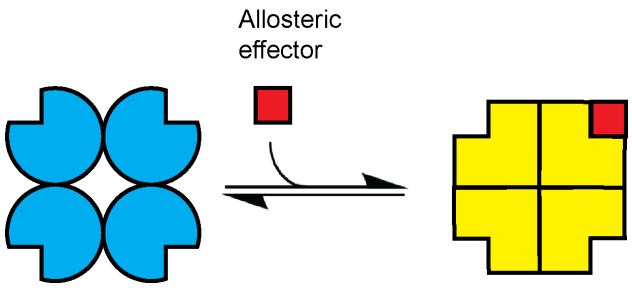


Figure 1.17: Cartoon representation of the Monod-Wyman-Changeux (MWC) model of allosteric transitions. A symmetric multimeric protein can exist in either of two conformational states, relaxed (R) or tense (T), that differ in their binding affinities for the ligand. Each subunit has distinct binding sites for the allosteric effector and the substrate (active site).

Discovery of Allosteric Regulation

Allosteric regulation was recognised well before the process was understood. The two earliest examples were the observation of a sigmoidal binding curve of haemoglobin to oxygen[86] (known as the Bohr effect), and the recognition of the activation of glycogen phosphorylase by AMP.[87] Although equations were derived to describe the processes,[88] it was not until the 1950s, with the discovery that Trp inhibited the synthesis of Trp synthetase in *E. coli*,[89] that the hypothesis of feedback regulation by pathway end products was formulated. This theory was demonstrated by Umbarger who showed that in many pathways an early enzyme is constructed to be strongly and specifically inhibited by the metabolic end product of the pathway.[90] The sites of regulatory molecule binding were renamed allosteric (Greek: allos "other", steric "site") to denote that the inhibitory effect occurs at a site distinct from the active site of the enzyme, and that the inhibitor is not simply a steric analogue of the substrate.[91–93] It has been suggested that allosteric sites may be more useful drug targets than active sites despite the difficulty in predicting the chemical nature of effectors and consequent slow discovery.[91]

In the past, allostery was typically referred to as "cooperativity", a term used to describe allosteric regulation in proteins composed of multiple subunits, where binding of a ligand to one subunit increases the affinity

of the other subunits for the same ligand (known as homotropic allostery). Examples were limited to multimeric proteins such as the classic haemoglobin and aspartate transcarbamoylase models.[85,94] Most allosteric effects can be explained by the concerted MWC model (Figure 1.17) or by the sequential Koshland-Nemethy-Filmer (KNF) model (Figure 1.18).

The central postulate of the MWC model is the existence of two distinct symmetric states of a protein composed of multiple subunits, R (relaxed) and T (tense). All subunits must exist in the same state. An equilibrium is assumed to exist between these two states of a protein and the position of this equilibrium is affected by ligand (or substrate) binding. Although the ligand can bind either protein state, the R form has a higher binding affinity for the ligand resulting in a shift of the equilibrium position towards the R form, in turn providing more high affinity binding sites for the ligand. This model explains the characteristic sigmoidal binding curve observed for proteins which display cooperative transitions between two distinct states, as a small ligand concentration will lead to a vast increase in the proportion of molecules in the R state and therefore will lead to a high association of ligand to the protein. However, the MWC model does not account for negative cooperativity.

The KNF model or sequential model for allosteric transitions (Figure 1.18) proposes that a conformational change in one subunit of a protein does not induce a similar change in the other subunits therefore all enzyme subunits do not necessitate the same conformation. Binding of a ligand induces a conformational change in the subunit to which it binds. There is no requirement for conservation of symmetry, the subunits can assume different conformations. If the interactions between the subunits are tightly coupled, binding of the ligand could cause the adjacent subunit(s) to adopt a conformation having a greater or lesser affinity for the ligand.

Today, allostery is considered to be a far more general property of proteins and it is also recognised in monomeric proteins (Figure 1.16).[96–98]

Catalysis and allostery of proteins arise, in principle, from an intrinsic

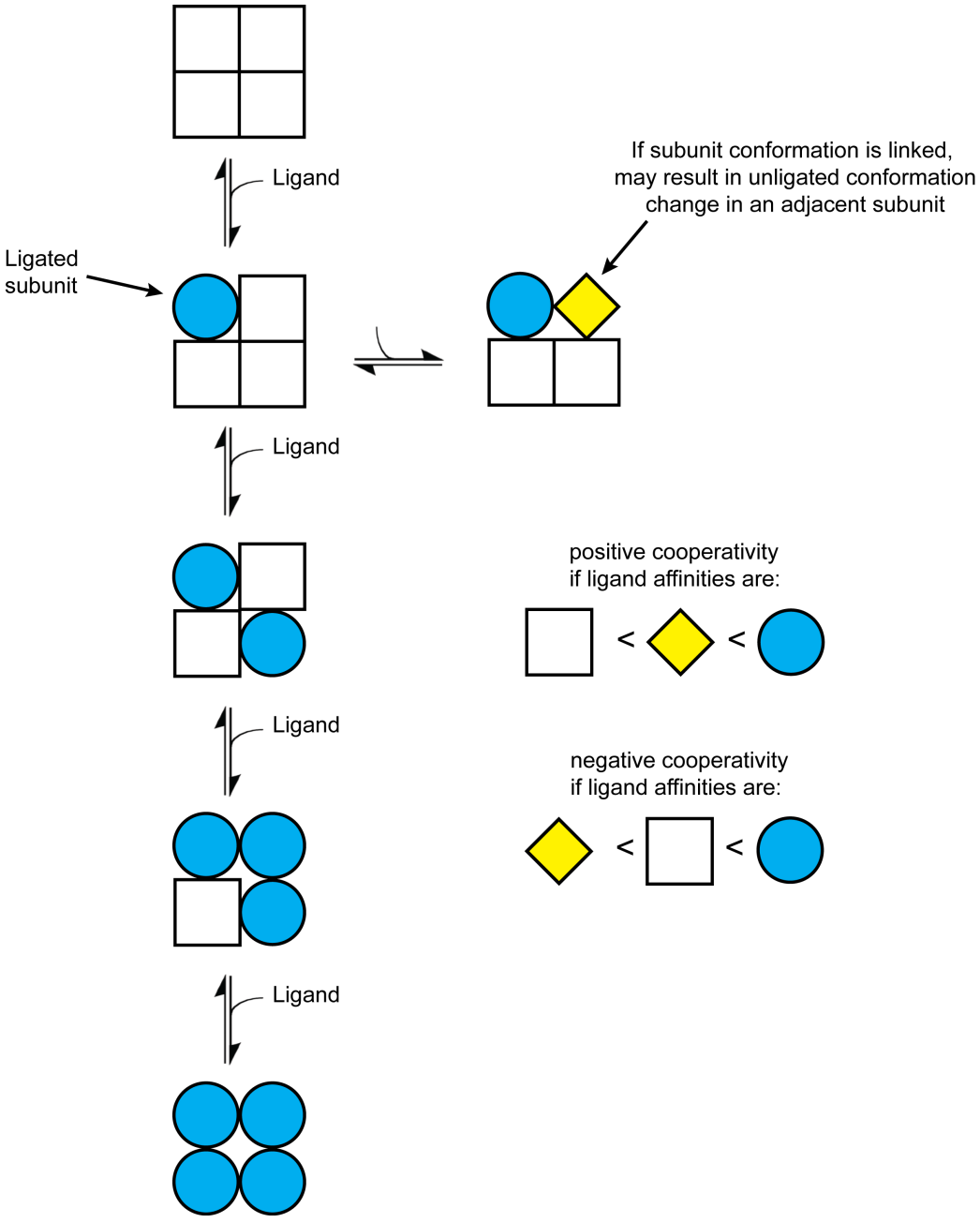


Figure 1.18: Schematic representation of the conformational states of a tetrameric protein according to the Koshland-Nemethy-Filmer (KNF) model for allosteric transitions. Illustration adapted from Garrett and Grisham.[95]

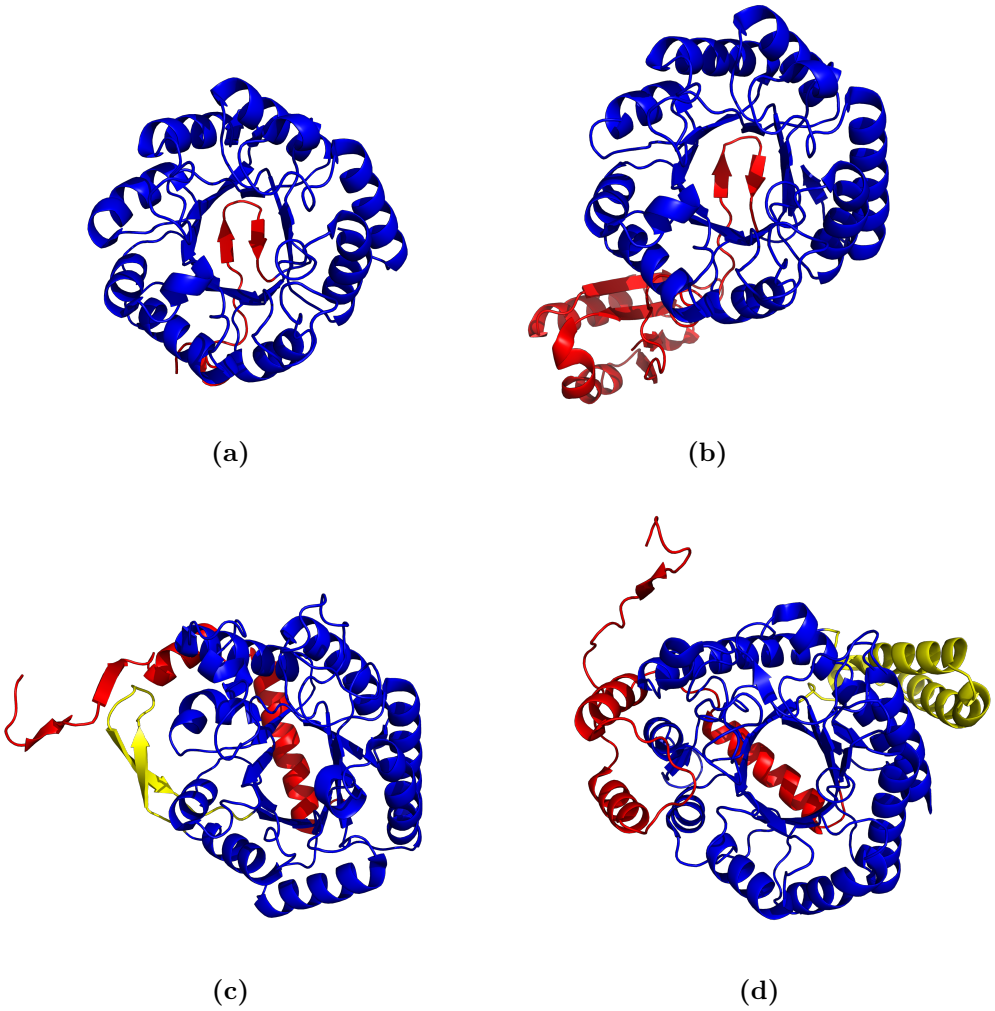
property of proteins, their conformational flexibility. This property allows the protein to interconvert between populations of conformers. The binding of a ligand can perturb this equilibrium and change the population distribution and/or the conversion rates between conformers. This results in modified properties of the active site or of binding-site geometries, and thereby rates of catalysis or binding affinities.[99,100]

The concept of large conformational changes being associated with remote ligand binding has been debated, as significant backbone deformations are *not* always a defining characteristic of allosteric effects. Changes in protein dynamics may also contribute to allostery.[101,102] Stably-folded proteins undergo fluctuations, especially of amino acid side chains, and it has been suggested that this imparts conformational entropy which could be important for regulation.[101]

Allosteric regulation is an efficient means for proteins to sense the changing conditions of the cell and respond to maintain cellular homeostasis. It provides a mechanism to directly sense the levels of small molecules, and can be very rapid (on the same order as the rate of diffusion) and readily reversible. Unlike covalent protein modification, no cellular energy is expended in allosteric signalling.

1.3.4 Regulation of DAH7PS

Typical of many biosynthetic pathways, the overall flux through the shikimate pathway is controlled via allosteric regulation of DAH7PS activity by pathway intermediates or end products. ^{13}C-NMR studies using whole cells of *E. coli* have shown that feedback inhibition of DAH7PS is the main method for controlling carbon flux through the shikimate pathway and cellular levels of aromatic compounds in microorganisms.[20] Different species appear to have adopted a variety of control patterns for DAH7PS activity and these divergent mechanisms of control are associated with the different types of DAH7PS (Figure 1.19). Some organisms have two or more isozymes, which are feedback inhibited by end products or intermediates of the pathway, or a combination



(a) **(b)**

(c) **(d)**

Figure 1.19: Different strategies for the feedback regulation of the DAH7PS enzymes (a) *Pfu*DAH7PS, (b) *Tma*DAH7PS, (c) *Eco*DAH7PS and (d) *Mtu*DAH7PS. Barrel regions are coloured blue, and (putative) regulatory regions are coloured red (N-terminal) and internal extensions are coloured yellow.

of end products. DAH7PSs from some organisms are not regulated.[40,48,103]

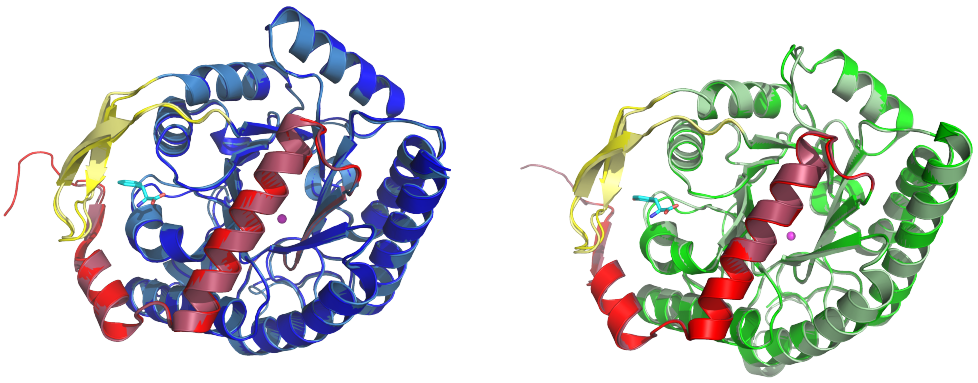
Type Iα DAH7PSs

One approach to allostery adopted by the Iα DAH7PSs is mediated through the expression of multiple isozymes. Scattered throughout the *E. coli* genome are genes encoding for three isozymes of DAH7PS, designated DAH7PS(Tyr), DAH7PS(Phe) and DAH7PS(Trp), and defined by the amino acid to which they show sensitivity. These proteins are encoded by the genes *aro*F, *aro*G and *aro*H, respectively.[5] Each isozyme is feedback inhibited by their associated aromatic amino acid.

The majority of the DAH7PS activity in *E. coli* arises from the *Eco*-DAH7PS(Phe) isozyme (eighty percent), with twenty percent of the total DAH7PS activity through *Eco*DAH7PS(Tyr), and the *Eco*DAH7PS(Trp) enzyme contributing less than one percent.[81] Similarly, *Neurospora crassa* and *Salmonella typhimurium* express three DAH7PS isozymes. However, the genome of *S. cerevisiae* encodes for only two isozymes, which have been shown to be sensitive to only Phe or Tyr.[33]

Crystal structures of both *Eco*DAH7PS(Phe) and *Sce*DAH7PS(Tyr) show the regulatory amino acids, Phe and Tyr, bind to these enzymes in very similar positions. The N-terminal extension and an internal extension between helix α5 and β6 together create a single binding site on the N-terminal end of the $(\beta/\alpha)_8$ barrel to which the inhibitor molecule binds, resulting in inhibition of the enzyme (Figure 1.20).[32–34,104]

When the ligand-bound structures of either *Eco*DAH7PS(Phe) or *Sce*-DAH7PS(Tyr) are overlaid on their unliganded forms, only very minor differences are observed in the structures (Figure 1.20) with only a slight perturbation of the binding loops responsible for substrate binding. It has been proposed for *Eco*DAH7PS(Phe) that the inhibitory signal is propagated from the allosteric site and affects the active site in two ways. A cascade of conformational changes are thought to result in perturbation to the interaction of



(a) (b)

Figure 1.20: Monomer subunits of type Iα showing limited conformational change on ligand binding. (a) Overlay of the monomer of *Eco*DAH7PS(Phe) with (PDB code 1KFL) and without (PDB code 1QR7) Phe (cyan sticks) bound. (b) Overlay of Tyr bound and unbound *Sce*DAH7PS(Tyr) (PDB codes 1OF6 and 1OFB, respectively). N-terminal extensions are coloured red and internal extensions are coloured yellow. Bound structures are coloured in the corresponding, but lighter shade to the unbound structures. The active-site metal ion is shown as a magenta sphere.

DAH7PS with both substrates. These changes result in the loss of E4P binding and PEP binding in a non-catalytically competent flipped orientation.[104]

Sequence analysis of type Iα DAH7PSs reveals a highly conserved serine (Ser) residue in all Phe-regulated enzymes (position 219 *Sce*DAH7PS(Phe) numbering), and a highly conserved glycine residue in all Tyr-sensitive enzymes (position 226 *Sce*DAH7PS(Tyr) numbering).[33] Analysis of the crystal structure of *Sce*DAH7PS(Tyr) shows G226 is located at the bottom of the putative effector binding cavity. This cavity is significantly reduced in *Sce*-DAH7PS(Phe), caused by the side chain of the highly conserved Ser residue of *Sce*DAH7PS(Phe) protruding out into the cavity, instead of the simple hydrogen side chain of glycine (Gly) found in *Sce*DAH7PS(Tyr). A single amino acid substitution in this position from Gly to Ser (Figure 1.21) was shown to completely interchange sensitivity of these *S. cerevisiae* enzymes from Phe to Tyr, without effecting the kinetic properties of the enzyme.[33]

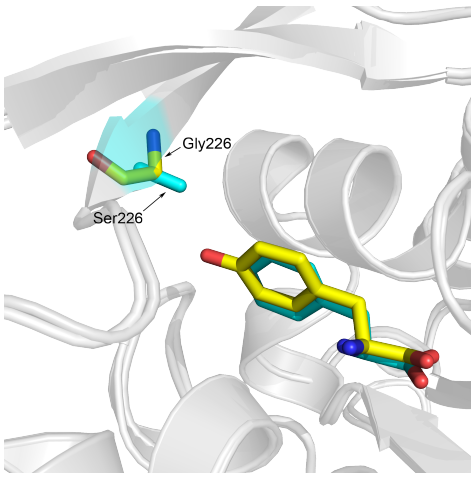


Figure 1.21: Overlay of *Sce*DAH7PS(Tyr) G226S mutant in complex with Phe (PDB code 1GO0) (unpublished) and wild-type *Sce*DAH7PS(Tyr) in complex with Tyr (PDB code 1OF6). Both the Tyr molecule and Gly226 of *Sce*DAH7PS(Tyr) are coloured yellow. The Phe molecule and residue Ser226 of the G226S structure are coloured cyan. Note: the sidechain of S226 is only modelled to the Cβ atom in all subunits of the G226S mutant structure.

Type Iβ DAH7PSs

Three members of the type Iβ DAH7PS subfamily have been structurally and functionally characterised.[38–40,48,64] Of these, two, *Pfu*DAH7PS[48] and *Ape*DAH7PS,[40] are constructed of only a simple core catalytic barrel (Figure 1.22a and 1.22b) and no allosteric effectors have been identified for these proteins. The other, *Tma*DAH7PS, has a FL domain attached to the N terminus of the barrel[39] (Figure 1.22c) and is inhibited by two of the three aromatic amino acids, namely Tyr and Phe.[64]

A single DAH7PS is found for *B. subtilis* with a chorismate mutase (CM) domain fused to the N terminus.[105] Similarly, *P. gingivalis* also has a CM fusion, but this extension is located at the C-terminal end of the protein.[105] In both of these cases CM is the regulatory domain, which facilitates feedback regulation by an intermediate of the pathway, chorismate, rather than pathway end products.[23,105]

Truncation of the terminal CM domains from *Bsu*DAH7PS and *P.*

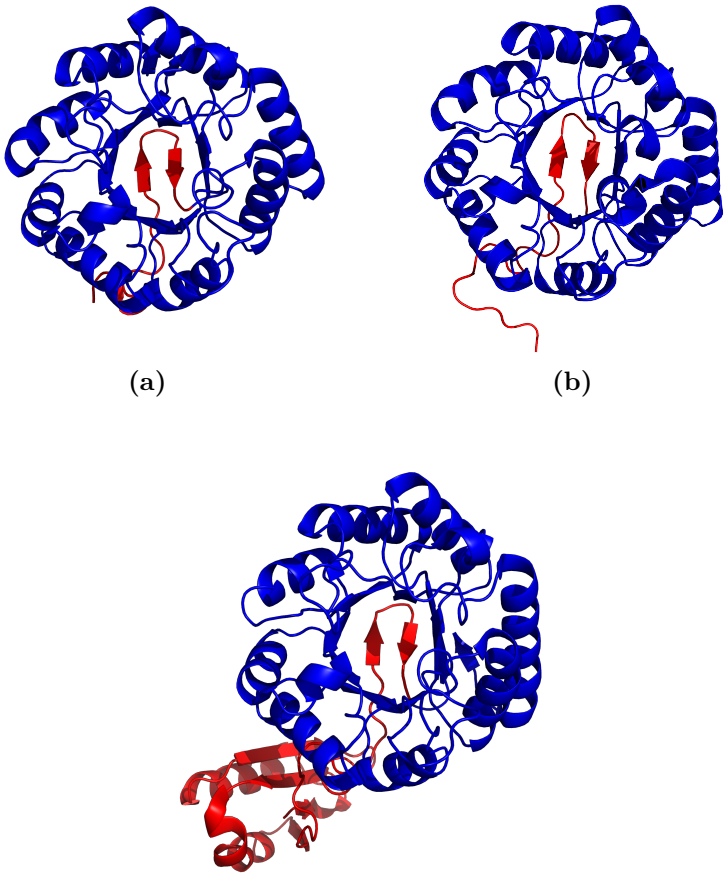


Figure 1.22: Structural comparison of type Iβ DAH7PS enzymes. (a) *Pfu*DAH7PS, (b) *Ape*DAH7PS and (c) *Tma*DAH7PS. The core barrel is coloured blue with the N-terminal extension coloured red.

gingivalis DAH7PS (*Pgi*DAH7PS) results in the loss of feedback inhibition by chorismate, strengthening the observation that feedback regulation is achieved through the extra domain elements. This suggests that these enzymes recruited the CM domain for the purpose of feedback regulation rather than to create bifunctional enzymes.[23]

Sequence analysis suggests that the Iβ subfamily also contains examples of DAH7PS isozymes, including those from *Thermus thermophilus*, where one appears to have a similar extension to *Tma*DAH7PS, and the other resembles

a CM-fused DAH7PS.

Although the N- or C-terminal extensions have been implicated in the feedback inhibition of the type Iβ DAH7PS, no mechanism has been proposed for the propagation of the regulatory signal from the allosteric site to the active site of the enzyme.[39] Furthermore, as no crystal structure of a ligand-bound type Iβ DAH7PS had been solved at the commencement of this thesis, the binding site for the ligands had not yet been determined.

Type II DAH7PSs

M. tuberculosis expresses a single type II DAH7PS enzyme and it has been recently discovered that *Mtu*DAH7PS displays a more complex method of regulation than most other DAH7PSs. *Mtu*DAH7PS is only weakly regulated by single aromatic amino acids, but displays greatest inhibition by a combination of two amino acids binding simultaneously to two binding sites formed by two distinct additions to the catalytic barrel.[66] This unusual but highly specialised mechanism of regulation, synergistic inhibition, allows the control of enzyme activity in response to the production of multiple pathway end products.

It is suggested that inhibition of *Mtu*DAH7PS catalysis occurs physically by perturbation of the loops associated with E4P binding.[66] Comparison of ligand-bound and unliganded crystal structures of *Mtu*DAH7PS show only minor conformational differences (Figure 1.23) and the loop containing the conserved KPRS motif, responsible for E4P binding, is slightly displaced but significantly more disordered on ligand binding, disfavouring binding of E4P. These results suggest that contributions from altered dynamics of the enzyme may be responsible for allosteric regulation.[106]

*Mtu*DAH7PS is not the only example of a synergistically regulated type II DAH7PS. Similarily, DAH7PS from *Corynebacterium glutamicum*, is reported to be strongly inhibited by a combination of Tyr and Phe, and shows further inhibition when Trp is added.[107]

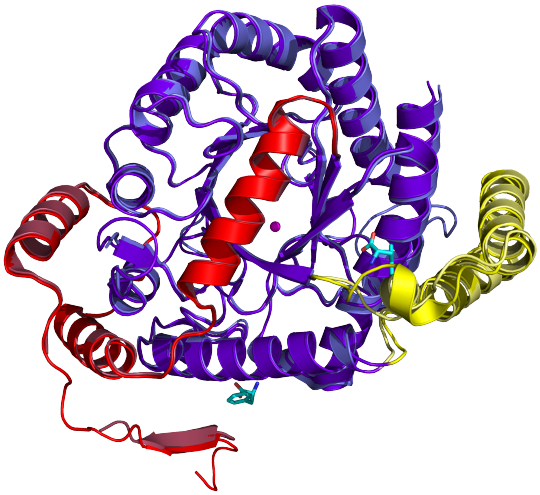


Figure 1.23: Superposition of an non-liganded (PDB code 3B70) and ligand-bound (PDB code 3KGF) monomer of *Mtu*DAH7PS. The non-liganded structure is coloured purple with the N-terminal extension coloured red and the internal barrel extension coloured yellow. The ligand-bound structure is shown in lighter shades of the same colours. The ligands, Phe and Trp, are shown as sticks and coloured cyan, and the active-site metal ion is shown as a magenta sphere.

As well as catalysing the first reaction on the pathway, *Mtu*DAH7PS is also involved in the regulation of the branch point that occurs at chorismate. *Mtu*DAH7PS forms a complex with *M. tuberculosis* CM.[108] This interaction results in significant activation of CM and the end products of this branch point, Phe and Tyr, inhibit this enhanced activity. It is unknown how CM activity is altered but the observation of stimulation of CM activity by association with DAH7PS is consistent with observations made for the same enzymes from *A. methanolica*.[43,109]

Although sequence analysis suggests some of the type II DAH7PS enzymes characterised to date appear to have the extensions required to facilitate regulation, allosteric effectors have not been identified for all type II enzymes.[47] Those with known inhibitors are subject to feedback regulation by either intermediates[42] or end products of the shikimate pathway.[103,110] Sequence alignments have also identified a subgroup of type II enzymes with a significant deletion (approximately 45 residues). It has been suggested

this deletion is a characteristic of type II DAH7PS enzymes involved in secondary metabolism[47] and that this region dictates allosteric specificity towards chorismate and Trp.[42]

1.4 Objectives of Thesis

The overall aim of this research is to unravel the evolutionary events that have led to complex and diverse contemporary mechanisms of regulation of the important biosynthetic enzyme DAH7PS. The prediction that the diverse regulation strategies for controlling DAH7PS activity have evolved from domain recruitment, by adding regulatory units to the catalytic barrel (Figure 1.24), will be investigated. To directly test this hypothesis, chimeric proteins with combinations of catalytic and regulatory domains from the type Iβ DAH7PSs will be constructed and characterised to ascertain whether regulation can indeed be introduced and eliminated as proposed. The aim is to investigate whether the putative FL regulatory domain of the type Iβ DAH7PS enzymes does control allostery and to uncover how this regulatory domain (if in fact it is) modulates enzyme activity. This information will be used to better understand more general regulation strategies of DAH7PS.

Specifically, the following studies are described in this thesis:

- the truncation of *Tma*DAH7PS by removing the putative regulatory domain to ascertain whether the extra domain element is essential for feedback regulation, catalytic activity and stability of the enzyme;
- an investigation of how the regulatory domain of *Tma*DAH7PS facilitates and communicates feedback inhibition to the active site;
- a study of the interactions between the putative regulatory domain and TIM-barrel domains, which are essential to transmit the feedback regulation signal to the active site;
- mutagenesis experiments designed to uncover the interactions between the regulatory molecule and ligand that govern the specificity and

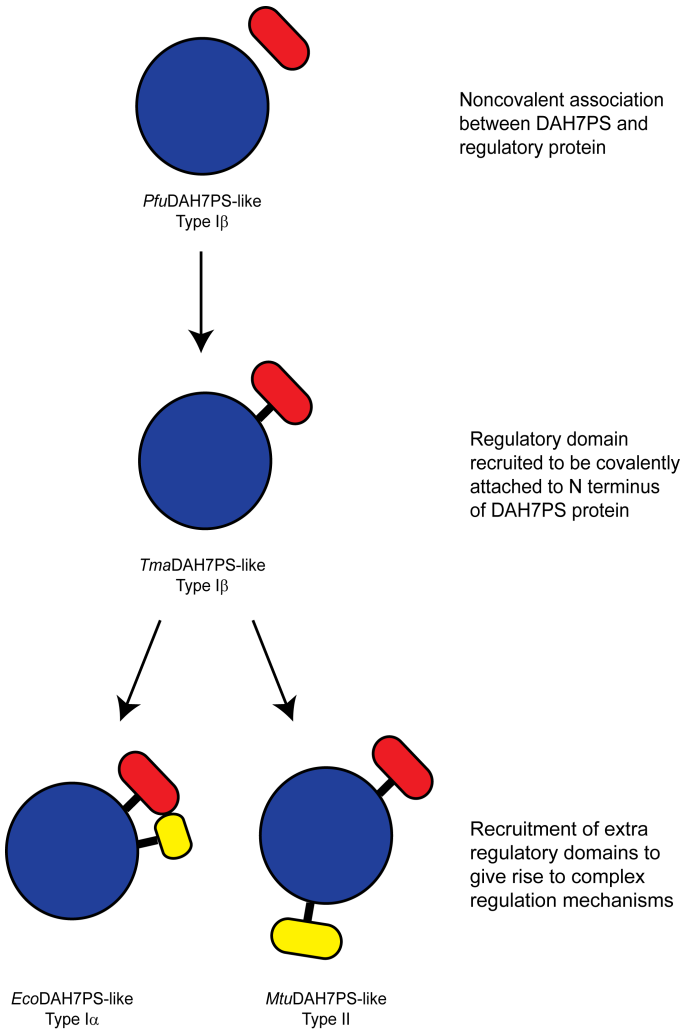


Figure 1.24: Proposal for the evolution of complex regulatory systems in DAH7PSs

effectiveness of the inhibitory molecule;

- investigations into whether the single amino acid exchange responsible for switching sensitivity between Phe and Tyr in the type Iα DAH7PS can be applied to type Iβ DAH7PS enzymes;
- the construction of a chimeric protein composed of a putative regulatory domain from a regulated DAH7PS with the TIM-barrel from an unregulated DAH7PS, to uncover whether allosteric regulation can be introduced by simple fusion of these two domains.

Chapter 2

Expression and Biochemical Characterisation of Type Iβ *Thermotoga maritima* DAH7PS

2.1 Introduction

Hyperthermophiles are a unique group of microorganisms that have a remarkable ability to grow at a temperature of 70 ℃ or above and have been isolated from geothermally-heated environments including deep-sea hydrothermal vents.[111] The hyperthermophiles are highly divergent microorganisms phylogentically, and although most of the hyperthermophilic genera are classified as Archaea, two groups, Thermotogales and Aquificales, belong under the domain Bacteria.[111,112] *T. maritima*, from the order Thermotogales, was originally isolated from geothermally-heated marine sediment at Vulcano, Italy and grows at temperatures between 55 ℃ and 90 ℃ with an optimal growing temperature of 80 ℃.[113] Along with *Aquifex* species, *T. maritima* is thought to be one of the most thermophilic Eubacteria currently known. Evolutionary

studies have placed this bacterium on one of the slowest evolving branches of the domain Bacteria[114,115] and it has therefore been suggested that study of this bacterium may provide insight into the divergence of the classes of the DAH7PS enzyme.[64] The genome of *T. maritima* encodes for a single DAH7PS that belongs to the type Iβ family. Although the characterisation of this enzyme had already been reported at the time of this study,[64] it was thought that additional characterisation was important to serve as a comparison for further studies. The expression, purification, biochemical characterisation and inhibition of *Tma*DAH7PS is presented in this chapter.

2.2 Cloning and Expression

Prior to the start of these studies, Dr. Mark Patchett had used standard methodologies to amplify the *Tma*DAH7PS gene (locus tag TM0343) from purified *T. maritima* MSB8 (DSM 3109) genomic DNA, and ligated it into the pT7-7 vector.[116] Plasmid DNA was isolated from ampicillin (Amp)-resistant colonies of transformed *E. coli* XL1-Blue cells, sequenced, and transformed into *E. coli* BL21(DE3)-Rosetta cells (Novagen) (for further details refer to Chapter 8).

The expression of *Tma*DAH7PS produced protein of the predicted subunit molecular weight and the *Tma*DAH7PS insert in the expression plasmid was sequenced and found to be identical to the desired sequence. The cells were grown and protein expression induced by the addition of isopropyl-1-thiol-β-D-galactopyranoside (IPTG). Cells were harvested four hours after induction, lysed by sonication on ice and cell debris removed by centrifugation. The soluble fraction was found to exhibit DAH7PS activity.

Attempts to replicate the reported method for expression and purification of *Tma*DAH7PS were made but resulted in very small yields of protein.[64] Numerous strategies were employed to improve the expression level and yield of soluble protein. Methods trialled included the use of auto-induction media and varying lysis conditions. Detergent-based methods and different temper-

atures and duration of sonication were tested, none of which appeared to make any difference to the overall yield of soluble protein within the crude lysate.

The gene encoding *Tma*DAH7PS was inserted into the alternative expression vector, pET151-D-TOPO® (Invitrogen). pET151D® allows expression of recombinant protein containing a N-terminal 6×His fusion tag for ease of detection and purification. The N-terminal tag also includes a Tobacco Etch Virus (TEV) protease cleavage site to enable removal of the tag using TEV after protein purification. Numerous attempts at purification of this construct were made, however the majority of protein was lost during the immobilised metal affinity chromatography (IMAC) step, so the pT7-7 construct was used for non-tagged expression.

Occasionally, sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) of the protein expressed from the pT7-7 vector revealed the presence of a slightly smaller molecular weight protein under the band for *Tma*DAH7PS. This smaller protein band was attributed to the cleaving of a small number of residues from a terminal end of *Tma*DAH7PS by native *E. coli* proteases present in the crude lysate. The addition of protease inhibitors, prior to lysis, appeared to decrease the appearance of this double banding. Although the yield of soluble protein was not ideal, expression from the pT7-7 construct was used for purification.

2.3 Purification

A three step purification protocol was developed based on the heat stability and calculated isoelectric point (pI) (6.3) of *Tma*DAH7PS. Fractions were selected and pooled on the basis of purity rather than for total enzyme recovery. Whole cells were lysed on ice using a cell disruptor, and the resulting lysate divided into small aliquots, heat treated, left to cool to room temperature, and clarified by centrifugation. Following heat treatment, SDS-PAGE analysis showed a number of contaminants were still present in the sample and hence,

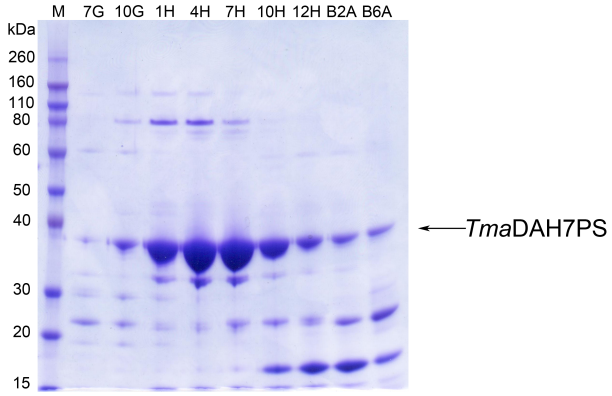


Figure 2.1: SDS-PAGE analysis of *Tma*DAH7PS purification by SEC. Lane M: molecular weight marker; other lanes correspond to the fraction loaded from the SEC column.

further purification steps were required.

Hydrophobic-interaction chromatography (HIC) is a chromatography technique that separates biomolecules according to their hydrophobic properties. Although most hydrophobic amino acids are buried in the interior of globular proteins, some of them are exposed on the protein surface. Proteins are separated due to the strength of the interaction between these exposed residues with hydrophobic groups attached to an uncharged matrix. Proteins bind at high ionic strength and are eluted from the column as the ionic strength is lowered. This technique was employed to further purify *Tma*-DAH7PS from contaminating proteins. Ammonium sulfate was added to the supernatant to a final concentration of 1 M, before the supernatant was applied to a hydrophobic interaction column equilibrated in a high salt buffer (1 M $(NH_4)_2SO_4$). When a gradient from 1 M to 0 M ammonium sulfate was applied, the protein was found to elute across the entire gradient. Subsequent purifications therefore used a stepwise elution with an immediate change from high salt to low salt. Fractions corresponding to the 280 nm absorbance peak observed at zero percent salt were pooled and concentrated.

Size-exclusion chromatography (SEC) was employed as a final polishing step for all purifications. Interestingly, the protein peak eluted from the SEC column was often found to be contaminated by a smaller molecular weight

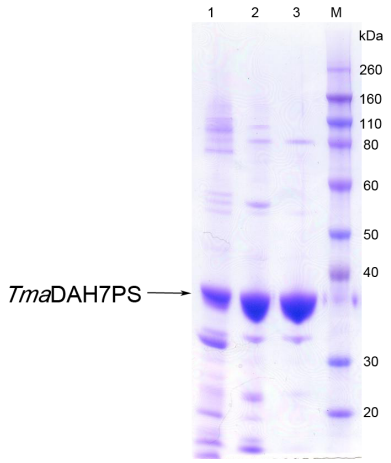


Figure 2.2: SDS-PAGE analysis of the purification of *Tma*DAH7PS. Lane 1: supernatant after heat treatment; lane 2: pooled HIC fractions; lane 3: pooled SEC fractions; lane M: molecular weight marker.

protein (Figure 2.1). Maximum separation of *Tma*DAH7PS and the smaller contaminating protein was achieved by slow elution from the column (Figure 2.2).

Due to small yields of protein (<2 mg L^{-1} of culture volume), alternative purification methods were trialled in an attempt to improve protein yield. Purification consisting of heat treatment, followed by two passes down an anion-exchange chromatography (AEC) was trialled and compared to the above method. Unfortunately this protocol resulted in a substantial loss of protein during the AEC and co-elution with other proteins resulted in a more contaminated sample.

In order to obtain a useful amount of purified protein (10 mg), 4 L growth cultures were necessary.

2.4 Confirmation of Molecular Weight

The molecular mass of *Tma*DAH7PS was 37382.6 ± 0.4 Da as determined by electrospray ionisation mass spectrometry. This is in close agreement with

the 37378.1 Da molecular weight calculated from the 338 amino acid open reading frame (ORF) encoding for *Tma*DAH7PS.

2.5 Kinetic Parameters

The steady-state kinetic parameters for freshly purified *Tma*DAH7PS were measured by following the disappearance of PEP absorbance at 232 nm. Although the optimal temperature for enzymatic activity of *Tma*DAH7PS has been reported to be 90 ℃,[64] the marked decrease in E4P stability over time at elevated temperatures restricts the temperature at which a continuous spectrophotometric assay can be performed. To minimise this effect, *Tma*-DAH7PS was assayed in the same manner described for *Pfu*DAH7PS,[38] by continuous spectrophotometric assay performed at 60 ℃ and initiated with E4P. An advantage of continuous assays over discontinuous assays is increased accuracy, which results from the measurement of a larger number of data points.[38] This assay system differs from that used in the characterisation of other type Iβ DAH7PS enzymes. The kinetic properties of *Pgi*DAH7PS,[23] *Bsu*DAH7PS,[23] *Ape*DAH7PS[40] and *Tma*DAH7PS[64] were explored by a discontinuous colourimetric or aminoff periodate-thiobarbituric acid assay system, or a combination of both continuous and discontinuous methods. All assays were initiated by the addition of E4P rather than enzyme.

Initial velocity rates of *Tma*DAH7PS catalytic activity were calculated

Table 2.1: Kinetic parameters for type Iβ DAH7PS enzymes

Enzyme	K_m^{PEP} (µM)	K_m^{E4P} (µM)	k_{cat} (s^{-1})	Temp. (℃)	Reference
*Pfu*DAH7PS	120 ± 20	28 ± 4	1.5 ± 0.1	60	Schofield[38]
*Tma*DAH7PS	12 ± 2	138 ± 11	5.0 ± 0.3	60	Wu[64]
*Pgi*DAH7PS	2724 ± 256	3638 ± 375	1.6 ± 0.1	37	Wu[105]
*Bsu*DAH7PS	142 ± 9	1400 ± 60	4.2 ± 0.2	37	Wu[105]
*Ape*DAH7PS	890 ± 30	280 ± 20	1.0 ± 0.2	60	Zhou[40]
*Tma*DAH7PS	4.9 ± 0.4	13 ± 1	11.7 ± 0.2	60	This study

Table 2.2: Optimal growth and assay temperatures for type Iβ DAH7PS enzymes

Organism/ DAH7PS source	Assay Temperature (°C)	Optimal Growth Temperature (°C)
P. furiosus	60	100[117]
T. maritima	60	80[113]
P. gingivalis	37	37[118]
B. subtilis	37	28–30[119]
A. pernix	60	90–95[120]

with the concentration of one substrate fixed while the other substrate concentration was varied. The apparent K_{m} values for PEP and E4P were found to be $4.9 \pm 0.4\,\mu\mathrm{M}$ and $13 \pm 1\,\mu\mathrm{M}$ respectively, and the k_{cat} value was calculated as $11.7 \pm 0.2\,\mathrm{s}^{-1}$.

It is difficult to accurately compare kinetic parameters between the type Iβ DAH7PSs as there is a large variation in assay conditions and the temperatures used are not always those that give maximal rates (Table 2.2). However, the K_{m} values determined for *Tma*DAH7PS are broadly in line with those previously reported for *Tma*DAH7PS (Table 2.1)[64] with the notable exception of the order of magnitude lower $K_{\mathrm{m}}^{\mathrm{E4P}}$ value.

2.6 Metal Activation

Consistent with other DAH7PS enzymes characterised to date, *Tma*DAH7PS requires a divalent metal ion for catalysis. Buffer and substrate solutions used in metal reactivation studies were treated with Chelex® (an ion-exchange resin that removes polyvalent metal ions) prior to use to avoid contamination from other metal sources. Enzyme activity was not detectable after the addition of $10\,\mu\mathrm{M}$ EDTA to the buffer. These results differ from other studies in which treatment of *X. campestris*, *S. aurantiaca* and *Streptomyces caespitosus* DAH7PSs with EDTA resulted in only partial deactivation of enzyme activity.[42,110,121]

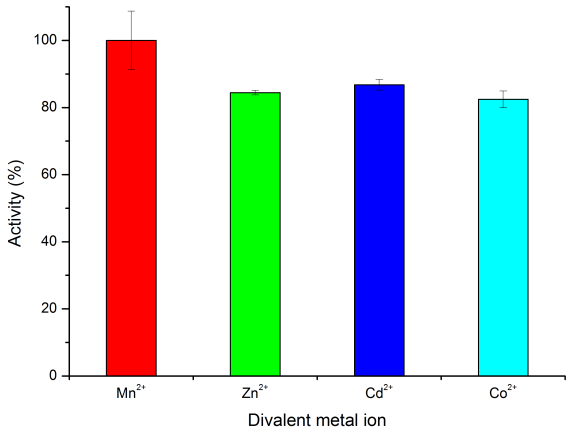


Figure 2.3: Extent of reactivation of *Tma*DAH7PS by various divalent metal ions.

Activity of the apo-enzyme could be restored by addition of Zn^{2+}, Cd^{2+} and Co^{2+}, with Mn^{2+} resulting in the fastest reaction rates (Figure 2.3). These results are consistent with previous reports for *Tma*DAH7PS and other type Iβ DAH7PS enzymes.[23,38,40,64]

2.7 Feedback Inhibition

As previously reported, the catalytic activity of *Tma*DAH7PS is sensitive to the presence of aromatic amino acids.[64] The degree of sensitivity of *Tma*-DAH7PS towards Phe, Tyr and Trp was investigated using standard assay conditions. The three amino acids were tested at a number of concentrations and it was found that Tyr caused a substantial decrease in catalytic activity, whereas Phe was less inhibitory (Figure 2.4). No significant reduction in activity was observed in the presence of Trp. In the presence of increasing concentrations of Tyr, enzyme activity decreased to a constant eighteen percent residual activity.

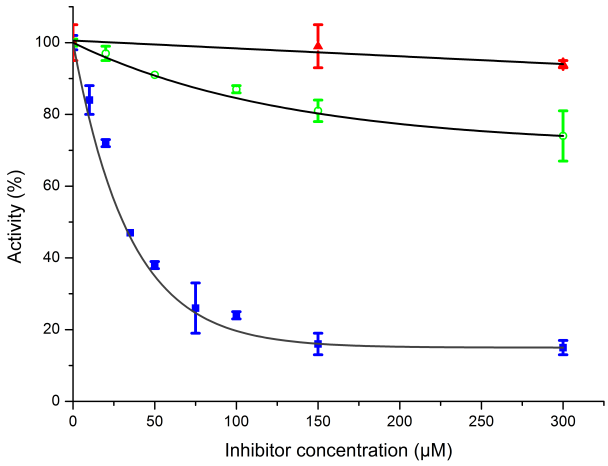


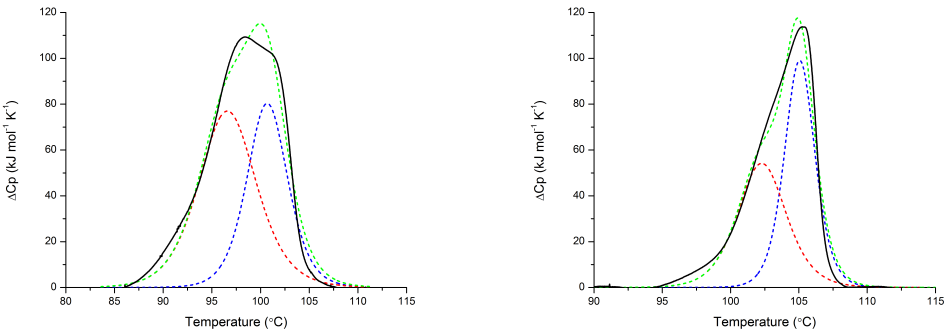
Figure 2.4: Response of wild-type *Tma*DAH7PS activity to increasing concentrations of Tyr (■), Phe (○) and Trp (▲). The data points were measured in triplicate with the error bars indicating the standard deviation.

2.8 Kinetic Parameters in the Presence of Tyr

Determination of the kinetic parameters of *Tma*DAH7PS in the presence of Tyr (150 µM) indicated that the wild-type protein has a reduced affinity for both substrates (K_m^{PEP} of 21 ± 1 µM, K_m^{E4P} of 118 ± 10 µM) as well as significantly lower k_{cat} of 2.79 ± 0.08 s^{-1} when compared with the parameters determined for this enzyme in the absence of this inhibitor (Table 2.1).

2.9 Differential Scanning Calorimetry

To investigate the heat stability of *Tma*DAH7PS and to probe any changes to thermal stability in the presence of Tyr, differential scanning calorimetry (DSC) experiments were carried out to determine the protein denaturation temperature (T_m). The result confirmed the high thermal stability of *Tma*-



(a) **(b)**

Figure 2.5: Differential scanning calorimetry of wild-type *Tma*DAH7PS in the (a) absence and (b) presence of Tyr showing the subtracted raw data (—), first T_{m} (---), second T_{m} (---) and the overall fit of the two state model (---).

DAH7PS, which denatured in two transitions, with the first T_{m} occurring at 96.8 ℃ and the second at 101 ℃. Addition of Tyr to the enzyme resulted in increased heat stability, with two thermal transitions observed at 102 ℃ and 105 ℃ (Figure 2.5).

2.10 Discussion

*Tma*DAH7PS was successfully purified and has been shown to be a thermally stable, metal-dependent enzyme that is inactivated by Tyr. Although a different purification method was developed, metal dependency and inhibition results are in agreement with what has been previously reported.[64] Mn^{2+} was the most activating divalent metal ion for *Tma*DAH7PS, although the order of metal reactivation series differs. Tyr and Phe inhibit the catalytic activity of the enzyme with Tyr being the more potent inhibitor.

Although it is suggested the N-terminal domain of *Tma*DAH7PS is required for the allosteric regulation of the enzyme,[39] further investigation is required to affirm this hypothesis. Will removal of this domain result in a catalytically active, unregulated enzyme?

Chapter 3

Expression and Biochemical Characterisation of Truncated *Thermotoga maritima* DAH7PS

3.1 Introduction

Previous studies have suggested that the extra domain elements found fused to the catalytic barrel of the DAH7PSs were acquired for the purpose of feedback regulation.[32–34,39,104,122] It is considered likely that contemporary DAH7PSs evolved from a primitive, unregulated and undecorated type Iβ-like enzyme and, while some members have remained unregulated, others have acquired an extra domain for regulation by feedback inhibition. The type Iα and type II subfamilies have evolved in a more complex manner, acquiring insertions or extensions within the $(\beta/\alpha)_8$ barrel itself, in addition to those found at the N terminus, to function as regulatory elements.[105]

The type Iβ subfamily has been found to be the most widely distributed

type of DAH7PS in nature, and it has therefore been proposed that this group of contemporary enzymes are the most closely related to the ancestral DAH7PS.[42] This subfamily is therefore an ideal system for examining the acquisition of feedback regulation in these enzymes. Type Iβ DAH7PSs consist simply of the $(\beta/\alpha)_8$ barrel alone, such as the unregulated *Pfu*DAH7PS[48] and *Ape*DAH7PS[40] or have an N-terminal or C-terminal extension to the catalytic barrel. The Iβ enzymes also include the DAH7PS-CM fusion proteins.[105]

*Tma*DAH7PS is a member of the type Iβ subfamily. This enzyme has been shown to have a FL domain at its N terminus that has been implicated in the sensitivity of this enzyme towards the aromatic amino acids.[39,64] The expression using various vectors, *in vivo* solubilisation, purification and biochemical characterisation of *Tma*DAH7PS with its N-terminal FL-domain removed is described in this chapter.

3.2 Choice of Truncation Site

To examine the role of the putative regulatory extension of *Tma*DAH7PS, a truncated mutant lacking this N-terminal domain was developed ($^{\text{trunc}}$*Tma*-DAH7PS). The basis of the truncation was determined from a structural comparison with the unregulated *Pfu*DAH7PS enzyme (Figures 3.1 and 3.2). To correspond to the start of the *Pfu*DAH7PS protein $^{\text{trunc}}$*Tma*DAH7PS started at Leu71.

3.3 $^{\text{trunc}}$*Tma*DAH7PS – pET151D-TOPO®

3.3.1 Cloning and Expression

DNA corresponding to the ORF of $^{\text{trunc}}$*Tma*DAH7PS was amplified from the plasmid containing the gene encoding for wild-type *Tma*DAH7PS, purified, concentrated and ligated into the pET151/D-TOPO® vector using TOPO®

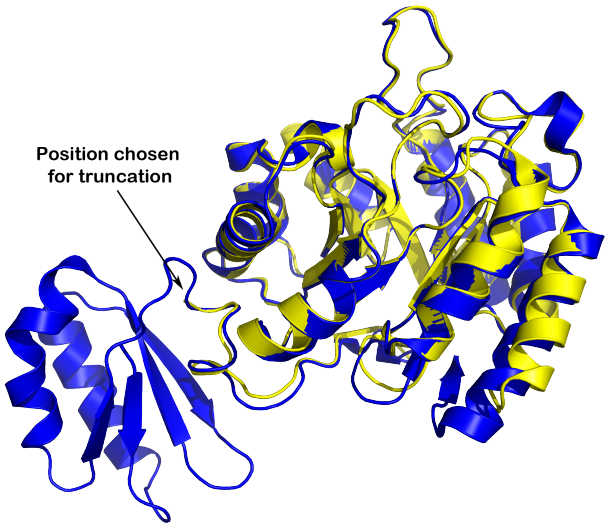


Figure 3.1: Overlay of the monomer units of the regulated *Tma*DAH7PS (PDB code 1RZM) coloured blue, with the unregulated, naturally truncated *Pfu*DAH7PS (PDB code 1ZCO), coloured yellow.

cloning. This vector enables the expression of a clone with a N-terminal TEV protease cleaveable His-tag, which after cleavage leaves the desired protein with six non-native amino acids at the N terminus. The $^{\text{trunc}}$*Tma*DAH7PS ORF containing plasmid was transformed into *E. coli* One Shot® BL21 Star™(DE3) cells and DNA sequencing of the plasmid confirmed the identity of the $^{\text{trunc}}$*Tma*DAH7PS ORF. Cells were grown, protein expression induced, cells lysed and the resulting crude lysate examined. SDS-PAGE analysis showed an over-expressed protein of the predicted molecular weight for $^{\text{trunc}}$*Tma*DAH7PS (approximately 30 kDa). Following cell lysis and centrifugation, equal amounts of the protein of interest were evident in the supernatant and cell debris. This distribution of protein was attributed to incomplete lysis of a small volume of resuspended cells. Preliminary activity assays confirmed the presence of DAH7PS catalytic activity.

Figure 3.2: ClustalW sequence alignment of type Iβ DAH7PS sequences. *Tma*DAH7PS has a N-terminal ferredoxin-like (FL) putative regulatory domain. Both *Pfu*DAH7PS and *Ape*DAH7PS are naturally occurring truncated DAH7PS enzymes while $^{\text{trunc}}$*Tma*DAH7PS has been engineered to resemble an undecorated DAH7PS enzyme. The Tobacco Etch Virus (TEV) protease recognition site of the $^{\text{trunc}}$*Tma*DAH7PS pET151/D-TOPO® construct is highlighted in yellow.

3.3.2 Purification

As $^{\text{trunc}}$*Tma*DAH7PS was expected to have similar physical properties to *Pfu*DAH7PS, a similar purification strategy was adopted.[38] Heat treatment was investigated as a potential one-step purification method. Enzymes tolerant of high temperatures can be easily purified by simple heat treatment after expression in a mesophile such as *E. coli*, as most of the host's own proteins denature and precipitate. This results in rapid and partial purification of thermostable enzymes, free from contaminating activities.[123] Heat treatment trials were conducted at 60 ℃, 70 ℃ and 80 ℃, for 30 min to 60 min. All conditions significantly purified the protein with no distinguishable difference in protein purity between temperatures or period of time. For subsequent purifications the supernatant post-lysis was heat treated at 60 ℃ for 30 min and the samples were left to cool to room temperature to ensure maximum precipitation of contaminating proteins.

IMAC, using a chelating-Sepharose column charged with either Ni^{2+} or Co^{2+}, was employed as a second purification step to take advantage of the polyvalent His-tag introduced to the N terminus of the protein. Unfortunately, only small amounts of protein were recovered from these columns and SDS-PAGE analysis of the flow through, urea suspensions, and regeneration of columns to assess affinity of the protein for the column, were unable to locate the majority of the protein.

The ability of TEV protease to cleave the polyvalent His-tag introduced to the N terminus of $^{\text{trunc}}$*Tma*DAH7PS was then assessed on protein that had been heat treated (Figure 3.3). Different concentrations of TEV, temperature and duration of incubation were trialled and all suggested $^{\text{trunc}}$*Tma*DAH7PS undergoes only partial cleavage of the His-tag by TEV. Due to the inability to utilise the His-tag in the purification of $^{\text{trunc}}$*Tma*DAH7PS, expression of the protein in this vector was abandoned.

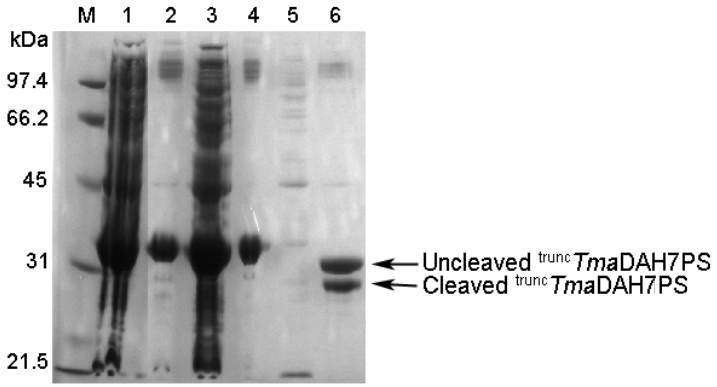


Figure 3.3: ^{trunc}*Tma*DAH7PS pET151/D-TOPO® SDS-PAGE analysis showing only partial TEV protease cleavage. Lane M: molecular weight marker; lane 1: resuspended cell pellet; lane 2: crude lysate after sonication; lane 3: resuspended cell debris after sonication; lane 4: supernatant post heat treatment; lane 5: resuspended debris after heat treatment; lane 8: supernatant after TEV protease treatment left overnight at 4 ℃.

3.4 ^{trunc}*Tma*DAH7PS – pDEST15 and pDEST566

3.4.1 Cloning and Expression

Given the unsuccessful attempts to purify ^{trunc}*Tma*DAH7PS expressed in the pET151/D-TOPO® vector, expression in an alternative vector was investigated. The gene encoding for ^{trunc}*Tma*DAH7PS was amplified using the pET151/D-TOPO® vector as a template and standard polymerase chain reaction (PCR) methodologies, and ligated into a Gateway®-adapted entry vector, pDONR™221, using Gateway® technology. This vector permits recombination cloning into multiple destination vectors allowing optimal expression of the gene of interest. Entry clones were transformed into One Shot® TOP10 cells. Colony PCR was used to check for colonies with gene inserts of the approximate size expected for ^{trunc}*Tma*DAH7PS and directionality of the insert.

Expression clones were then generated in two vectors encoding different N-terminal fusion tags, pDEST15 (glutathione S-transferase (GST)) and pDEST566 (maltose binding protein (MBP)), using Gateway® technology. A major advantage of GST and MBP tags is that they allow affinity purification of the fusion protein with mild elution conditions, and they are known to enhance the solubility of their fusion partners.[124] Following recombination, the vectors were transformed into *E. coli* One Shot® BL21 Star™ (DE3) and Rosetta™ 2 (DE3) cells. Rosetta™ 2 (DE3) host cells were used to try to improve the expression level of $^{\text{trunc}}$*Tma*DAH7PS. These cells are BL21 derivatives designed to enhance the expression of proteins that contain codons rarely used in *E. coli*. As the nucleotide sequence of $^{\text{trunc}}$*Tma*DAH7PS has rare codon usage, Rosetta™ 2 (DE3) cells might be expected to improve protein expression.

Trials were conducted to compare the lysis methods of sonication and a more gentle detergent-based lysis method in both the BL21 Star™ (DE3) and Rosetta™ 2 (DE3) cell lines. Following cell lysis and centrifugation, SDS-PAGE analysis showed the protein of interest was only evident in the cell debris of cells expressing the GST-tagged protein. However, the supernatant resulting from sonication of the BL21 Star™ (DE3) cells expressing the MBP-tagged protein contained some soluble protein (Figure 3.4), therefore this construct was pursued.

3.4.2 Purification

BL21 Star™ (DE3) transformants harbouring the pDEST566-MBP-$^{\text{trunc}}$*Tma*-DAH7PS plasmid were grown and protein expression was induced by addition of IPTG. Cell pellets were resuspended, sonicated and clarified by centrifugation.

The supernatant liquid was applied to a MBPTrap™ HP column (GE Healthcare) and $^{\text{trunc}}$*Tma*DAH7PS eluted with maltose. Fractions containing $^{\text{trunc}}$*Tma*DAH7PS were concentrated and desalted. The MBP-tag was cleaved from $^{\text{trunc}}$*Tma*DAH7PS using TEV protease and the protein was heat treated,

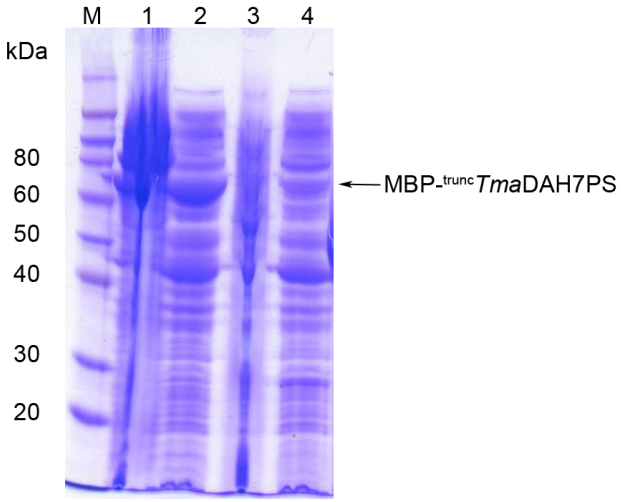


Figure 3.4: SDS-PAGE analysis of the expression of $^{\text{trunc}}$*Tma*DAH7PS from pDEST566 and pDEST15 vectors and lysis methodology. Lane M: molecular weight marker; lane 1: supernatant resulting from the sonication and lane 2: from detergent lysis of cells expressing the MBP-tagged protein. Lane 3 and 4: soluble fraction resulting from the sonication and detergent lysis of cells expressing the GST-tagged protein, respectively.

left to cool to room temperature, and centrifuged. The supernatant was reapplied to the MBPTrap™ in order to remove the MBP. $^{\text{trunc}}$*Tma*DAH7PS was eluted in the flow through and was further purified by SEC. Fractions from SEC containing $^{\text{trunc}}$*Tma*DAH7PS were pooled, concentrated and due to persistent contamination of MBP, reloaded to a MBPTrap™ to remove any remaining MBP-tag (Figure 3.5). Fractions containing $^{\text{trunc}}$*Tma*DAH7PS were pooled, concentrated, flash frozen and stored at −80 ℃.

3.4.3 Cleavage of the MBP-tag

Due to the incomplete cleavage of the His-tag observed for $^{\text{trunc}}$*Tma*DAH7PS expressed in the pET151/D-TOPO® vector, cleavage of the MBP-tag by TEV was assessed (Figure 3.6). Time trials revealed a majority of the cleavage occurred during the first hour, however only partial cleavage (approximately 50 percent) was ever achieved suggesting inaccessibility of the TEV recognition site or a high affinity association of MBP and $^{\text{trunc}}$*Tma*DAH7PS.

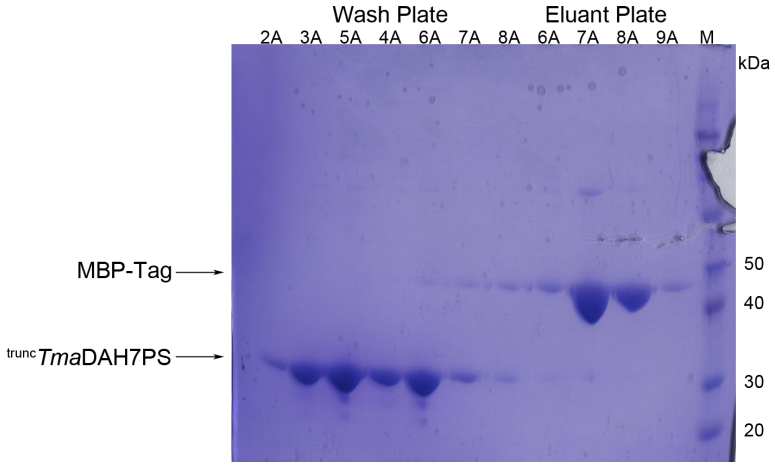


Figure 3.5: SDS-PAGE analysis of the final MBPTrap™ purification step of $^{\text{trunc}}$*Tma*DAH7PS. Pooled fractions of $^{\text{trunc}}$*Tma*DAH7PS collected from SEC were applied to a MBPTrap™ to try and eliminate the persistent contaminating cleaved MBP-tag. $^{\text{trunc}}$*Tma*DAH7PS eluted in the flow through of the wash plate and the MBP-tag in the eluant plate fractions.

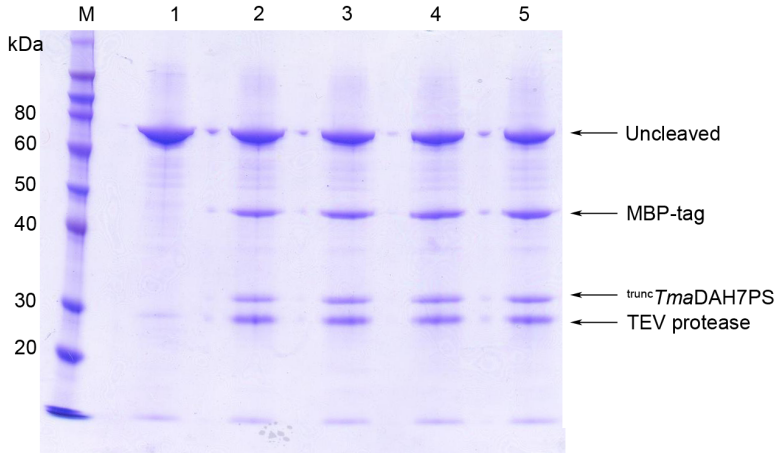


Figure 3.6: SDS-PAGE analysis of cleavage of MBP from $^{\text{trunc}}$*Tma*DAH7PS by TEV protease. Lane M: molecular weight marker; lane 1: $^{\text{trunc}}$*Tma*DAH7PS before incubation with TEV protease; lane 2: $^{\text{trunc}}$*Tma*DAH7PS after incubation with TEV for 1 h; lane 3: 2 h; lane 4: 3 h and lane 5: 4 h.

3.5 Confirmation of Molecular Weight

The molecular weight of $^{\text{trunc}}$*Tma*DAH7PS was 29885.9 ± 0.2 Da as determined by electrospray mass spectrometry. This is in close agreement with the 29880 Da molecular weight calculated from the 268 amino acid open reading frame coding for $^{\text{trunc}}$*Tma*DAH7PS and the three amino acid residues (Gly, Ser and Phe) left at the N terminus following cleavage of the MBP-tag.

3.6 Kinetic Parameters

The steady-state kinetic parameters for $^{\text{trunc}}$*Tma*DAH7PS were measured by following the disappearance of PEP absorbance, spectrophotometrically, at 232 nm and 60 °C. Initial velocity rates were calculated with the concentration of one substrate fixed while the other substrate concentration was varied. The apparent K_{m} values for E4P and PEP were 38 ± 3 µM and 6.6 ± 0.5 µM respectively, and the k_{cat} value was calculated as 18.2 ± 0.3 s^{-1}.

Compared to the full length wild-type protein, $^{\text{trunc}}$*Tma*DAH7PS has higher $K_{\text{m}}^{\text{E4P}}$ and $K_{\text{m}}^{\text{PEP}}$ values (refer to Chapter 2, Section 2.5). Intriguingly, this mutant protein was more active than wild-type *Tma*DAH7PS. The k_{cat} value for $^{\text{trunc}}$*Tma*DAH7PS is 1.6-fold higher than the k_{cat} observed for wild-type *Tma*DAH7PS (11.7 ± 0.2 s^{-1}).

3.7 Metal Activation

Consistent with wild-type *Tma*DAH7PS, $^{\text{trunc}}$*Tma*DAH7PS also requires a divalent metal ion for catalysis,[64] although there are some small differences in the activation by different metal ions when compared to the activation of the wild-type enzyme. To avoid metal contamination, substrate solutions used in metal reactivation studies were pretreated with Chelex® resin, an exchange resin that removes polyvalent metal ions, prior to use. Addition

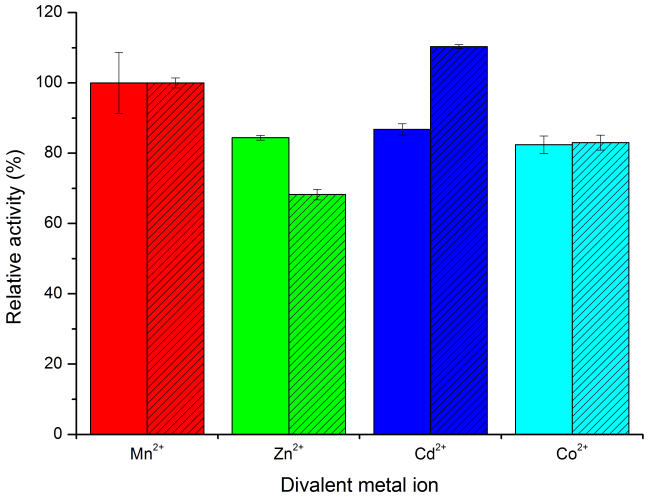


Figure 3.7: Extent of reactivation of $^{\text{trunc}}$*Tma*DAH7PS (patterned bars) by various divalent metal ions in comparison with wild-type *Tma*DAH7PS (plain bars). Error bars represent the standard deviation.

of 10 µM EDTA to the buffer abolished enzyme activity. Catalytic activity of the apo-enzyme could be restored by addition of 100 µM Zn^{2+}, Mn^{2+}, Co^{2+}, or Cd^{2+}, with Cd^{2+} resulting in the fastest reaction rates (Figure 3.7). The most activating metal ion for wild-type *Tma*DAH7PS was Mn^{2+}. To enable direct comparison with wild-type, this metal ion was also used for the characterisation of $^{\text{trunc}}$*Tma*DAH7PS.

3.8 Feedback Inhibition

The degree of sensitivity of $^{\text{trunc}}$*Tma*DAH7PS towards Phe, Tyr and Trp was investigated. The inhibitory effect of the three amino acids was tested at increasing concentrations. In marked contrast to wild-type *Tma*DAH7PS, the presence of the aromatic amino acids had no effect on the catalytic activity of $^{\text{trunc}}$*Tma*DAH7PS (Figure 3.8).

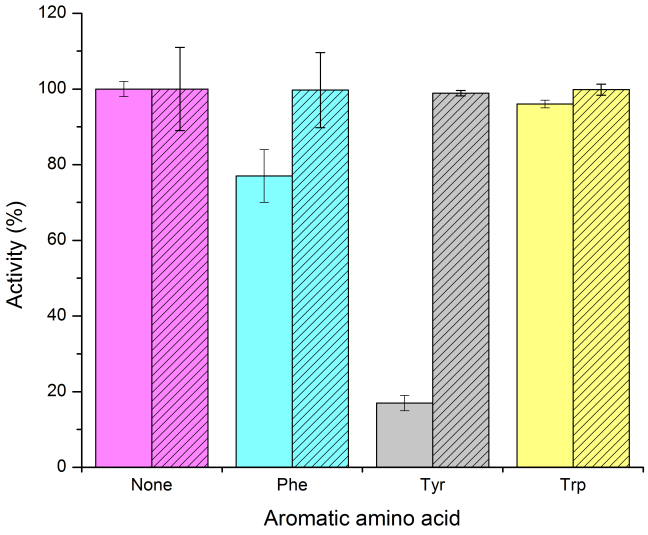


Figure 3.8: Activity remaining in the presence of the aromatic amino acids (300 µM) for $^{\text{trunc}}$*Tma*DAH7PS (patterned bars) in comparison to wild-type *Tma*DAH7PS (plain bars). Error bars represent the differences in measured values.

3.9 Crystallisation and Structure

3.9.1 Crystallisation Trials

The structure of $^{\text{trunc}}$*Tma*DAH7PS was sought to investigate any changes in protein structure compared to the wild-type enzyme. Crystallisation trials were performed at the Collaborative Crystallisation Centre (hosted by CSIRO) at 293 K. Initial screens were set up using the sitting-drop vapour diffusion method. The best diffracting crystal grew from a reservoir solution containing 8 % (w/v) PEG 4000, 0.1 M sodium acetate, pH 4.6 and 0.02 % (w/v) sodium azide (Figure 3.9).

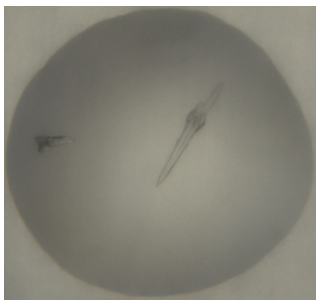


Figure 3.9: Crystals of $^{\text{trunc}}$*Tma*DAH7PS that provided the best diffraction.

3.9.2 Structure

$^{\text{trunc}}$*Tma*DAH7PS crystallised in the space group $P2_12_12_1$ (Table 3.1). The two monomers in the asymmetric unit show close structural alignment (RMSD of 0.15 Å[i]) and associate to form a dimer. The dimer interface, which buries approximately 13 percent of the solvent accessible surface area of each monomer, lies on an approximate two-fold rotation axis and includes helices α6, α7 and α8 and the β6-α6 loop.

The $^{\text{trunc}}$*Tma*DAH7PS dimer in the asymmetric unit is very similar to the equivalent dimer in the tetrameric wild-type enzyme (PDB code 1RZM, RMSD of 0.54 Å, Figure 3.10a and 3.10c). However, apart from the missing regulatory domain, the notable difference between $^{\text{trunc}}$*Tma*DAH7PS and *Tma*DAH7PS is the variation in quaternary structure in the crystal structure; $^{\text{trunc}}$*Tma*DAH7PS is dimeric, whereas the wild-type protein is tetrameric, with the tetramer formed by the interaction of helices (α_5, α_6, α_7 and α_8) and loops (β_5-α_5, β_6-α_6 and β_7-α_7) at the C-terminal end of the barrel.[39]

The $^{\text{trunc}}$*Tma*DAH7PS dimer overlays closely with the naturally truncated *Pfu*DAH7PS structure (PDB code 1ZCO, RMSD of 0.73 Å). Unlike the engineered $^{\text{trunc}}$*Tma*DAH7PS, *Pfu*DAH7PS is tetrameric in its crystal form, although it has been predicted to be dimeric in solution.[48] Native PAGE analysis of $^{\text{trunc}}$*Tma*DAH7PS under non-denaturing conditions indicates it adopts the same quaternary structure as *Pfu*DAH7PS (Figure 3.11).

[i] All RMSD values hereafter are calculated between equivalent Cα atoms.

Table 3.1: Crystal parameters, data collection, and refinement statistics for $^{\text{trunc}}$*Tma*DAH7PS

	$^{\text{trunc}}$*Tma*DAH7PS
Data collection	
Space group	$P2_12_12_1$
Cell dimensions	
a, *b*, *c* (Å)	53.5, 55.5, 166.9
α, β, γ (°)	90.0, 90.0, 90.0
Resolution range (Å)	38.57–2.00 (2.05–2.00)[a]
R_{merge}[b]	0.067 (0.293)
$I/\sigma(I)$	6.8 (2.0)
Completeness (%)	94.5 (96.3)
Redundancy	3.7 (3.7)
Refinement	
No. of reflections	120147
$R_{\text{work}}/R_{\text{free}}$[b]	0.155/0.209
No. of atoms	
Protein	4008
Ligand/ion	45
Water	238
B-factors (Å^2)	
Protein	29.5
Ligand/ion	50.4
Water	39.2
RMSD	
Bond lengths (Å)	0.006
Bond angles (°)	0.995
Ramachandran	
Most favoured (%)	98.4
Allowed (%)	1.0
Disallowed (%)	0.6
PDB code	3PG8

[a]Values in parentheses are for the highest resolution shell.
[b]R_{merge}= $\sum h \sum i \, |Ii(h) - \langle I(h) \rangle| \, / \sum h \sum i Ii(h)$ and is provided by SCALA.
$r = \sum \| F_{obs}| - |F_{cal} \| \, / \sum |F_{obs}|$, where F_{obs} and F_{cal} are the observed and calculated structure-factor amplitudes, respectively for the working set (R_{work}) and test set (R_{free}) of reflections and are provided by REFMAC5.

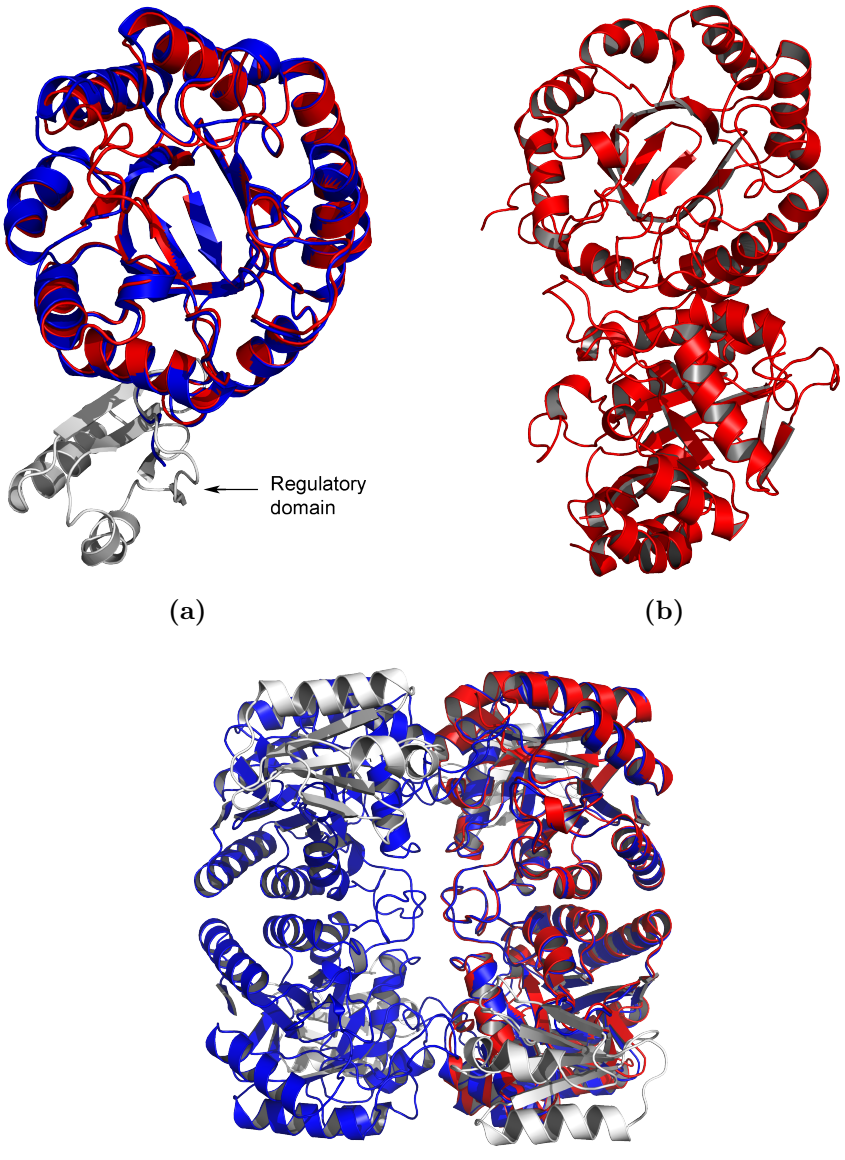


Figure 3.10: Structure of $^{\text{trunc}}$*Tma*DAH7PS. (a) $^{\text{trunc}}$*Tma*DAH7PS monomer (red) overlaid with wild-type *Tma*DAH7PS (PDB code 1RZM) monomer (blue with the regulatory domain in grey) (RMSD of 0.33 Å). (b) $^{\text{trunc}}$*Tma*DAH7PS dimer (red). (c) $^{\text{trunc}}$*Tma*DAH7PS dimer overlaid with *Tma*DAH7PS tetramer (blue with the regulatory domains in grey) (RMSD of 0.72 Å).

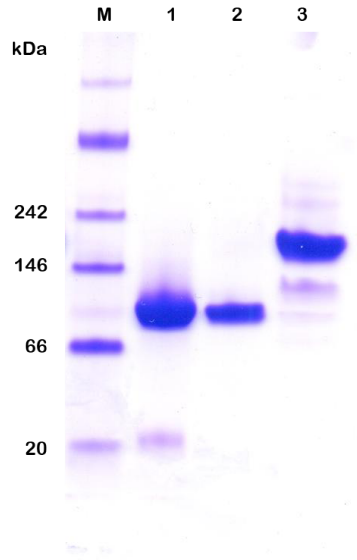
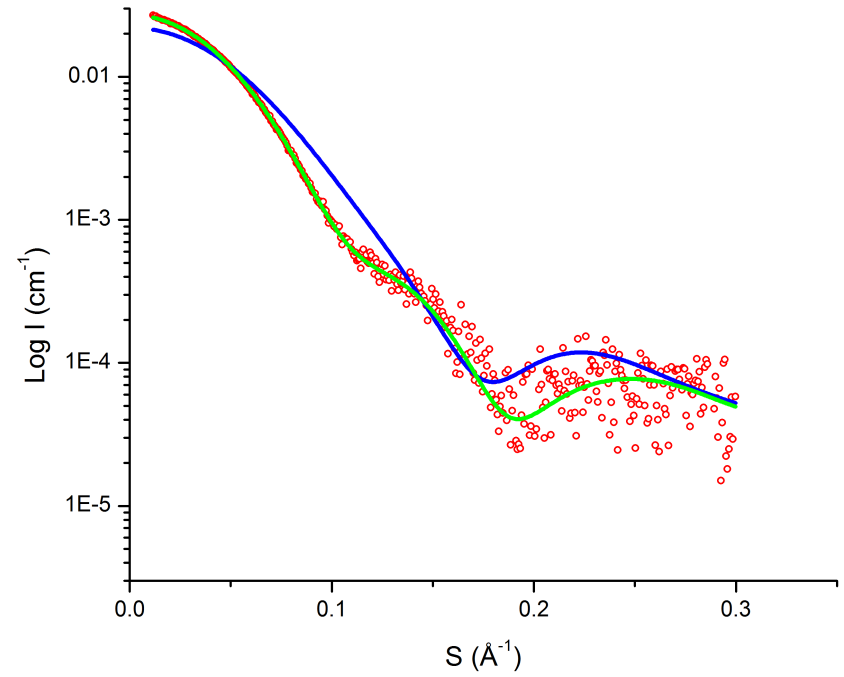


Figure 3.11: Non-denaturing PAGE analysis of $^{\text{trunc}}$*Tma*DAH7PS, *Pfu*DAH7PS and wild-type *Tma*DAH7PS. Lane M: molecular weight marker; lane 1: $^{\text{trunc}}$*Tma*DAH7PS; lane 2: *Pfu*DAH7PS; and lane 3: wild-type *Tma*DAH7PS.

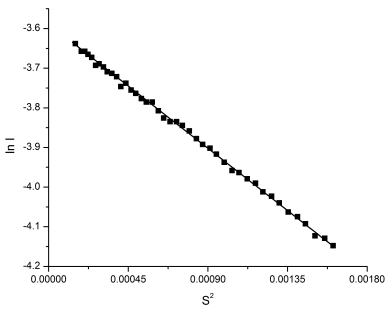
3.10 Small Angle X-ray Scattering

Small angle X-ray scattering (SAXS) was employed to further verify the quaternary structure of $^{\text{trunc}}$*Tma*DAH7PS in solution. For the scattering profile of $^{\text{trunc}}$*Tma*DAH7PS, the radius of gyration (R_{g}), as determined by Guinier analysis, was 32.2 ± 0.3 Å (Figure 3.12b). The pair distribution function ($P(r)$) was determined using the indirect Fourier transform method[125] and calculated the R_g to be 32.2 ± 0.2 Å and the maximum scattering particle (D_{max}) was approximately 90 Å (Figure 3.12c). This is similar for those values calculated for the tetrameric wild-type *Tma*DAH7PS crystal structure with the regulatory domain truncated from each of the monomers (truncation of residues 1–71 for all monomers), with a calculated D_{max} of approximately 91 Å. There is a notable difference when comparing these values to those calculated for the dimeric crystal structure of $^{\text{trunc}}$*Tma*DAH7PS (PDB code 3PG8, D_{max} value of approximately 85 Å).

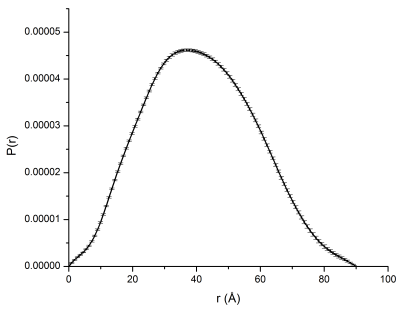
The differences are more apparent on direct comparison of the CRYSOL



(a)



(b)



(c)

Figure 3.12: SAXS measurements of $^{\text{trunc}}$*Tma*DAH7PS. (a) Theoretical scattering profiles were generated from crystallographic coordinates of both the dimeric $^{\text{trunc}}$*Tma*DAH7PS (—) and the tetrameric wild-type *Tma*DAH7PS with the regulatory domain truncated (residues 1–71) from all monomers (—), using CRYSOL. These theoretical plots were compared with the experimentally determined $^{\text{trunc}}$*Tma*DAH7PS data (○). (b) A Guinier plot was linear for $s \cdot R_g < 1.3$ for data collected and (c) $P(r)$ function calculated from $^{\text{trunc}}$*Tma*DAH7PS SAXS data was calculated using GNOM.[125] Uncertainty in $P(r)$ propagated from the $I(s)$ vs s profile is indicated by error bars.

fits (Figure 3.12a). The scattering data are a better fit with the theoretical profile calculated for a truncated tetramer of *Tma*DAH7PS (PDB code 1RZM), created by deletion of residues 1–71, than with the dimeric crystal form (PDB code 3PG8) when analysed with CRYSOL (χ^2_ν of 0.2 and χ^2_ν of 34.2, respectively). This clearly suggests that at the concentrations used in the SAXS analysis, $^{\text{trunc}}$*Tma*DAH7PS adopts a tetrameric quaternary structure in solution, which is in contrast to the result obtained from PAGE non-denaturing gel analysis (Figure 3.11). This suggests the truncated enzyme exists in a dimer-tetramer equilibrium, that has a $K_{\text{d}}^{4\rightarrow 2}$ significantly lower than the protein concentration used in SAXS experiment.

3.11 Discussion

Removal of the N-terminal domain of *Tma*DAH7PS generates a protein that is an efficient catalyst, yet insensitive to Phe and Tyr, conclusively demonstrating that this domain is required for the inhibition observed for the wild-type protein.[64] Furthermore, $^{\text{trunc}}$*Tma*DAH7PS has a higher catalytic efficiency than the wild-type enzyme. This phenomenon of higher catalytic activity by removal of putative regulatory domains has been noted for a number of enzymes including the dual specificity kinase, CLK1,[126] rice glutamate decarboxylase[127] and D-3-phosphoglycerate dehydrogenase,[128] suggesting that regulatory domains recruited by these enzymes act to conformationally constrain catalytic activity.[126,129] The increased catalytic activity observed for $^{\text{trunc}}$*Tma*DAH7PS implies that there is a catalytic penalty associated with the presence of the regulatory domain.

In crystal form, $^{\text{trunc}}$*Tma*DAH7PS is dimeric, unlike the tetrameric wild-type enzyme, suggesting that the extra domain helps maintain both quaternary structure and allosteric mechanism. However, different techniques of measurement suggest that tetrameric species of $^{\text{trunc}}$*Tma*DAH7PS are indeed possible. But the real question is what is the nature of the dimer-tetramer equilibrium and associated stabilities in physiologically relevant

conditions? This highlights the importance of highly accurate techniques such as analytical ultra centrifugation (AUC), and in particular sedimentation equilibrium AUC, for the determination and quantification of oligomeric species in equilibrium.

Interestingly, the dependence of quaternary structure on regulatory domains has been investigated for the first enzyme in the Ser biosynthetic pathway, D-3-phosphoglycerate dehydrogenase (PGDH).[128] PGDH is a homotetramer that is allosterically regulated by Ser. It was hypothesised that removal of the regulatory domain would create an enzyme mimic of the unregulated closely related dimeric enzymes, the D-2-hydroxyacid dehydrogenases. Truncation of the regulatory domain of PGDH eliminated feedback inhibition by Ser, but did not result in a change from tetrameric to dimeric quaternary structure. In mammals, Ser feedback has been replaced by transcriptional control yet sequence alignments from a variety of species show retention of the regulatory domain. From these observations it has been suggested that the regulatory domain of PGDH also functions to provide additional stability to quaternary structure. However from the observation of the dimer in the crystalline form of $^{\text{trunc}}$*Tma*DAH7PS, it appears that removal of the regulatory domain may be sufficient to disrupt quaternary structure.

Having confirmed that the extra N-terminal domain of the type Iβ *Tma*DAH7PS is essential for allosteric regulation, several key questions are evident. What is the mechanism used by this enzyme for regulation? Where does the inhibitor bind? And, how does ligand binding transmit the inhibitory signal?

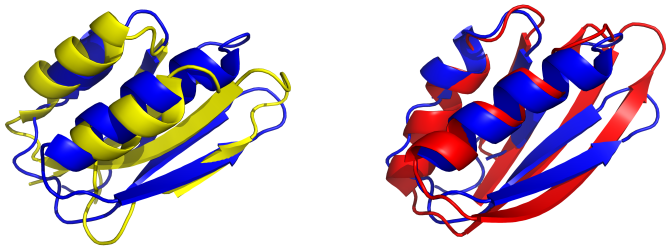
Chapter 4

Structural Characterisation of Type Iβ *Thermotoga maritima* DAH7PS

4.1 Introduction

Having established that the N-terminal domain of *Tma*DAH7PS is required for allosteric regulation, the key questions are how does this extra domain element foster the regulation of this enzyme, and where do the ligands bind? This chapter describes the crystallisation and crystal structure determination of the Tyr-bound *Tma*DAH7PS.

The structure of wild-type *Tma*DAH7PS had been previously determined, and the monomer consists of a catalytic barrel with an N-terminal extension, described as a FL domain.[39] However, this domain has a remarkable resemblance to an ACT domain (Figure 4.1), a structural motif found in enzymes responsible for amino acid biosynthesis, for the purpose of ligand binding.[130,131] This motif was first discovered in 1999,[132] when a search of a diverse set of proteins noted to be involved in amino acid or purine biosyn-



(a) **(b)**

Figure 4.1: Structural alignment of the N-terminal regulatory domain of *Tma*DAH7PS (blue) with the C-terminal regulatory domain of (a) ATP phosphoribosyltransferase (red) (PDB code 1NH8, RMSD of 4.59 Å for equivalent Cα atoms) and (b) the C-terminal regulatory domain of phosphoglycerate mutase (yellow) (PDB code 1PSD, RMSD of 4.48 Å).

thesis and regulated by specific amino acids, recognised this recurring motif. The proposed domain was named the ACT domain after the first letters of three of the proteins in which it was originally identified: aspartate kinase, chorismate mutase and tyrA (prephenate dehydrogenase). The first structure of an ACT domain was determined in 1995 with the crystal structure of *E. coli* D-3-phosphoglycerate dehydrogenase.[133]

4.2 Crystallisation Trials

As extensions to the core catalytic barrel of the DAH7PS enzymes have been shown to facilitate allosteric control of catalytic activity, it has been suggested that the N-terminal extension of *Tma*DAH7PS is required for allosteric control of this enzyme.[64] In order to shed light on the mechanism for allosteric regulation, attempts were made to crystallise *Tma*DAH7PS in the presence of its most potent inhibitor, Tyr, as no ligand-bound structure of the type Iβ DAH7PSs had been solved before these studies.

Crystallisation trials were performed at the Collaborative Crystallisation Centre (hosted by CSIRO) at 293 K. Diffracting crystals were obtained from

a reservoir solution containing 20 % (w/v) polyethylene glycol 3350, 0.2 M ammonium nitrate, and 0.02 % (w/v) sodium azide (Figure 4.2).

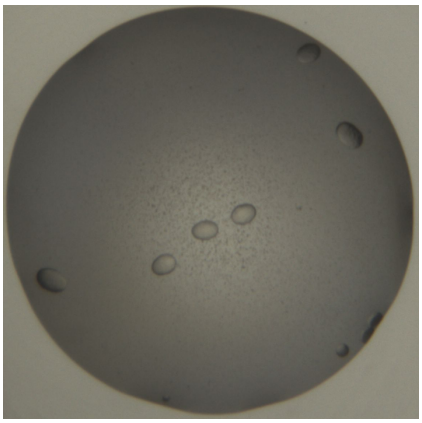


Figure 4.2: Crystals of *Tma*DAH7PS co-crystallised with Tyr

4.3 Structure of *Tma*DAH7PS in Complex with Tyr

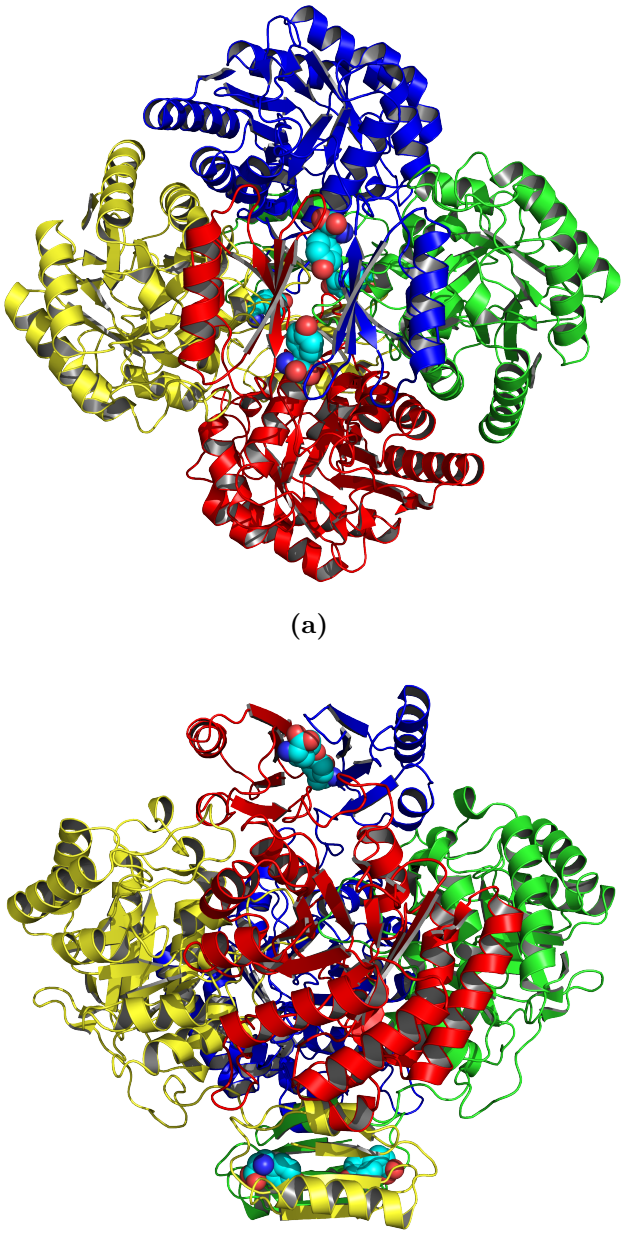
Diffraction data from a crystal in Figure 4.2 were collected at the Australian Synchrotron and solved by molecular replacement in the space group $P2_1$. The asymmetric unit contains eight monomers that display close structural alignment (maximum RMSD of 0.26 Å). The *Tma*DAH7PS monomer consists of an N-terminal regulatory domain with a βαββαβ fold (or ACT domain) arranged as a four-stranded antiparallel β-sheet with two α-helices (residues 1–63), tethered by a flexible linker region (residues 64–93) to the catalytic $(\beta/\alpha)_8$-barrel domain (residues 94–338).

The Tyr-bound structure is tetrameric (Figure 4.3) and the eight monomers in the asymmetric unit form two independent tetramers each of approximately 222 symmetry. These tetramers show close structural correspondence with respect to each other (RMSD of 0.43 Å). The loops extending from the C terminus of the barrel (β2-α2, β3-α3, β4-α4, β5-α5, and β6-α6), the α5 helix, and the C-terminal end of the α4 helix associate, burying approximately seven

Table 4.1: Crystal parameters, data collection, and refinement statistics for wild-type *Tma*DAH7PS-Tyr complex

	*Tma*DAH7PS-Tyr complex
Data collection	
Space group	$P2_1$
Cell dimensions	
a, b, c (Å)	81.6, 121.0, 133.4
α, β, γ (°)	90.0, 92.1, 90.0
Resolution range (Å)	52.57–2.35 (2.43–2.35)[a]
R_{merge}[b]	0.102 (0.528)
$I/\sigma(I)$	6.0 (2.0)
Completeness (%)	97.4 (98.6)
Redundancy	3.0 (3.0)
Refinement	
No. of reflections	104678
$R_{\text{work}}/R_{\text{free}}$[b]	0.171/0.222
No. of atoms	
Protein	20456
Ligand/ion	127
Water	565
B-factors (Å^2)	
Protein	40.9
Ligand/ion	30.4
Water	40.6
RMSD	
Bond lengths (Å)	0.003
Bond angles (°)	0.654
Ramachandran	
Most favoured (%)	97.7
Allowed (%)	1.8
Disallowed (%)	0.5
PDB code	3PG9

[a]Values in parentheses are for the highest resolution shell.
[b]$R_{merge} = \sum h \sum i \, |Ii(h) - \langle I(h) \rangle| \, / \sum h \sum i Ii(h)$ and is provided by SCALA.
$r = \sum \| F_{obs}| - |F_{cal} \| \, / \sum |F_{obs}|$, where F_{obs} and F_{cal} are the observed and calculated structure-factor amplitudes, respectively for the working set (R_{work}) and test set (R_{free}) of reflections and are provided by REFMAC5.



(a)

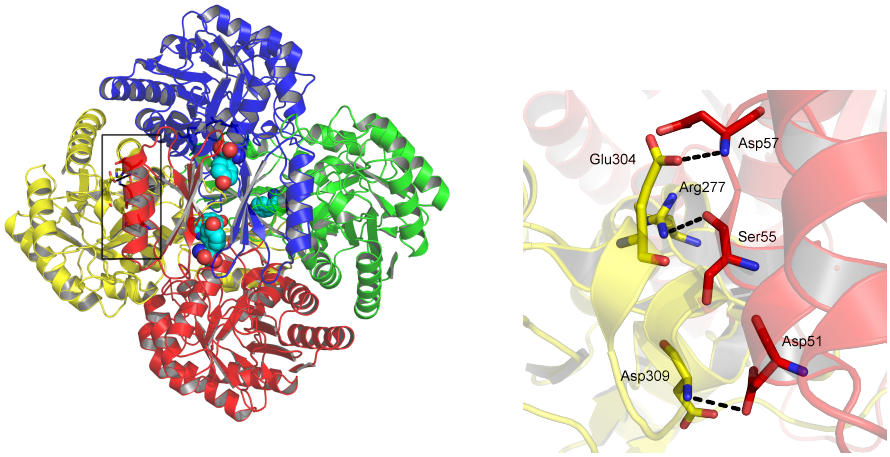
(b)

Figure 4.3: Tyr-bound structure of *Tma*DAH7PS (a) looking down on, and (b) side-on to the tetramer. Tyr molecules are shown as spheres, with the carbon atoms coloured cyan. Each chain has been assigned a different colour.

percent of the surface area of the barrel to form the tetramer. As is the case for all $(\beta/\alpha)_8$-barrel enzymes, the active site is located at the C-terminal end of the barrel. The arrangement of the monomers in the tetramer results in alternating active sites around the tetramer so that the diagonally opposite active sites face the same side of the tetramer plane.

In addition to the contacts between the barrels in the Tyr-bound tetramer, which correspond closely with those observed for the Tyr-free structure, there is significant interaction between the diagonally opposed N-terminal regulatory domains, where the anti-parallel β-sheets of diagonally opposite regulatory domains face each other in their altered position (Figure 4.3a). The regulatory domain and linker regions from diagonally opposite monomers interact with each other through hydrogen bonds between His29 and Ser31, Glu35 and Arg46, and Arg36 and Glu75.

The regulatory domain, in its altered position in the Tyr-bound structure, also makes contact with residues of the β8-α8 loop at the C-terminal end of the adjacent barrel, through hydrogen bonds between the main-chain NH of Asp51 to the $O^{\epsilon 2}$ of Asp309, carbonyl oxygen of Ser55 to $N^{\eta 1}$ of Arg277,



(a) **(b)**

Figure 4.4: Hydrogen-bonding interactions between the regulatory domain and C-terminal end of adjacent barrel in the Tyr-bound structure of *Tma*DAH7PS (a) looking down on the tetramer, and a (b) close-up view.

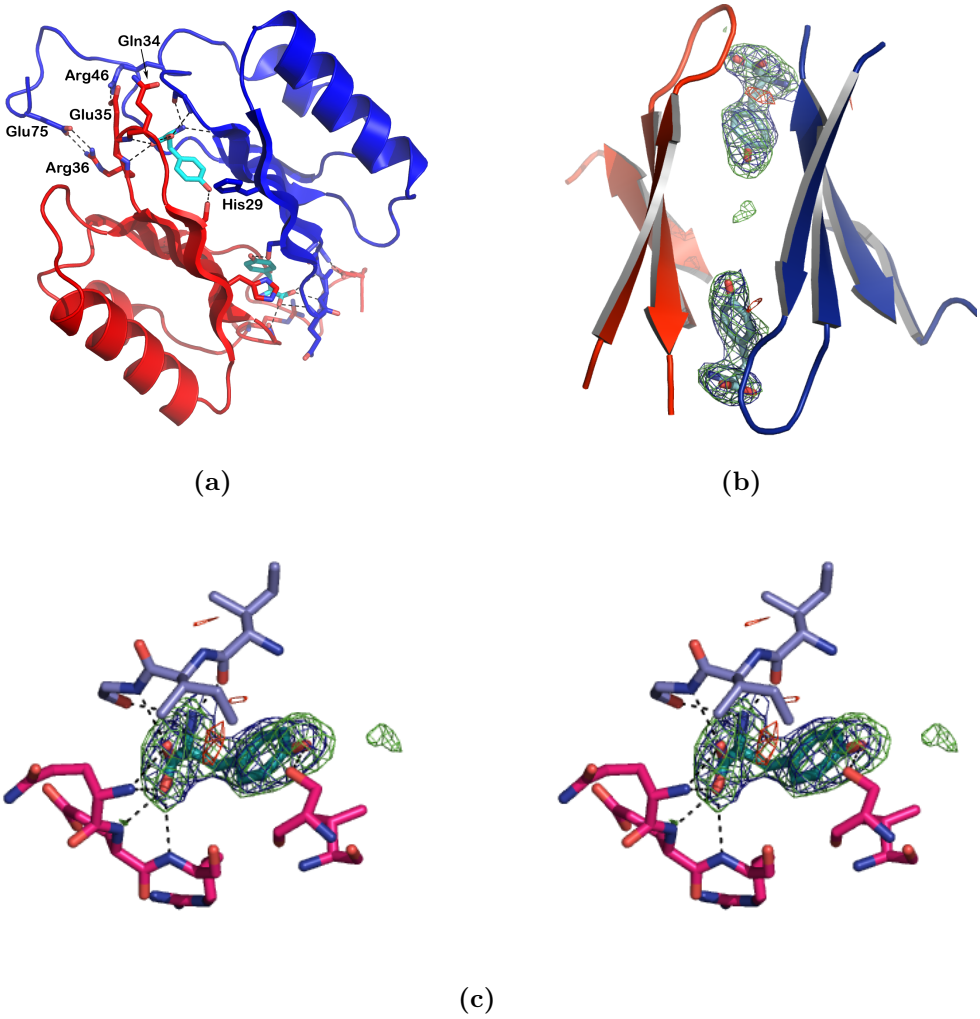


Figure 4.5: Structure of the Tyr binding site when in complex with *Tma*DAH7PS. (a) Interactions of regulatory domains of diagonally opposite molecules of the tetrameric *Tma*DAH7PS. Tyr is shown as cyan sticks. (b) Electron density ($2F_o-F_c$, contoured at 1σ(blue); F_o-F_c, contoured to 3σ (green); F_o-F_c, contoured to -3σ (red) showing Tyr binding between regulatory domains of opposing monomers and (c) stereo view. Residues contributed by the regulatory domains of different monomers are coloured either magenta or blue.

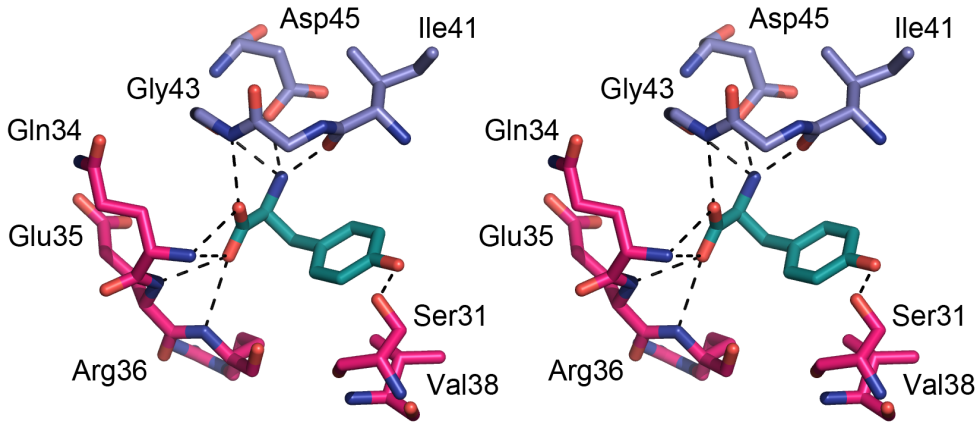


Figure 4.6: Stereo view of the Tyr binding site. Residues contributed by the regulatory domains of different monomers are coloured either magenta or blue. Hydrogen-bonding interactions are indicated by dashed lines.

Asp57 to the main-chain NH of Glu304, and for some subunits, between the side chain of Ser55 to the carbonyl oxygen of either Glu304 or Pro303 (Figure 4.4).

An omit map generated during structure refinement prior to adding the Tyr molecules to the model, clearly suggests that Tyr is bound. In addition, the refined electron density for the Tyr molecules was very well resolved (Figures 4.5b and 4.5c).

In addition to direct contacts, the two Tyr molecules provide further interactions between diagonally opposed regulatory domains by forming hydrogen bonds with residues of both domains (Figure 4.5 and 4.6).

Four Tyr molecules are bound per *Tma*DAH7PS tetramer. The carboxylate functionality of Tyr is bound by hydrogen-bonding interactions between the main-chain amide of residues Gln34, Glu35 and Arg36 (located in the $\beta 2_r$-$\beta 3_r$ loop of the regulatory domain) of the same monomer, and to the main-chain amide of Gly43 of the regulatory domain of the opposing monomer (Figures 4.5a and 4.6).

In addition, the amino functionality of the bound Tyr hydrogen bonds

with the main-chain carbonyl oxygens of both Ile41 and Gly43 and to Asp45 $O^{\delta 1}$. There are also several hydrophobic chains that approach within 4 Å of the Tyr phenyl ring, including Val38, and Ile42 and Val65 of the opposing monomer. The phenolic hydroxyl group of Tyr hydrogen bonds to Ser31 (distance of 2.6 Å), providing a molecular explanation for the increased potency of Tyr over Phe as an allosteric inhibitor.

4.3.1 Comparison of the Ligand-free and Tyr-bound *Tma*DAH7PS Crystal Structures

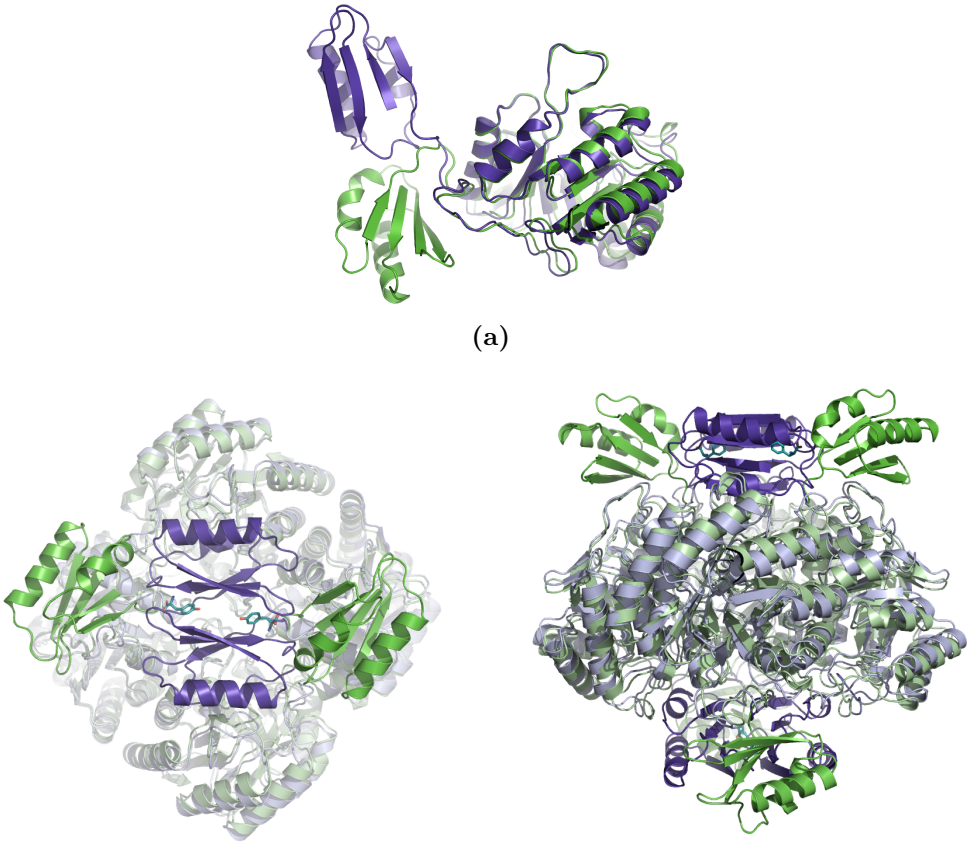
The striking difference between the Tyr-bound and ligand-free structures is in the position of the regulatory domain relative to the tetramer formed by the catalytic barrels, which is obvious when the structures are overlaid (Figure 4.7a).

In the Tyr-free structure, the regulatory domains from different monomers do not interact. In the presence of Tyr, these domains are repositioned via a large rotation (approximately 114°). Specifically, residues Val65, Leu66 and Lys67 are responsible for the deviation of the main chain in the flexible linker region (Table 4.2).

The consequence of this large displacement of the regulatory domains is that in the Tyr-bound structure, the regulatory domains effectively cap the C-terminal ends of the barrels that house the active sites (Figure 4.7b). This

Table 4.2: Phi (ϕ) and psi (ψ) angles of the residues in the flexible linker region of *Tma*DAH7PS in the absence and presence of Tyr

Residue	Non Tyr-bound		Tyr-bound	
	ϕ	ψ	ϕ	ψ
Val65	−100	97	−106	−52
Leu66	−63	20	−80	147
Lys67	−126	153	−71	141



(a) **(b)** **(c)**

Figure 4.7: Comparison of Tyr-bound (coloured purple) and Tyr-free (coloured green) *Tma*DAH7PS structures revealing major domain reorganisation on Tyr binding. (a) An overlay of the Tyr-bound and Tyr-free monomer. An overlay of Tyr-bound and Tyr-free tetramer looking (b) down on and (c) side on, at the tetramer.

closed structure occludes the active site of each barrel to substrates, thereby inhibiting the enzyme. In the absence of Tyr, the protein adopts an open conformation, and substrates can access the active site.

A comparison of the catalytic barrel regions of liganded and unliganded *Tma*DAH7PS reveals little overall difference between the two structures (RMSD of barrel domains is 0.27 Å). However, there are some changes in the loops on the C-terminal end of the barrel that contribute to substrate binding. On Tyr binding the β2-α2 loop, responsible for the binding of E4P is

displaced relative to its position in the Tyr-free structure and becomes poorly defined. This suggests that there may be greater conformational flexibility in this part of the active site when Tyr is bound to the regulatory domain. This is consistent with kinetic data, which suggests E4P binding is weakened in the presence of Tyr.

Intriguingly, the closed form of the enzyme observed in the Tyr-bound structure is also observed in a structure of *Tma*DAH7PS deposited in the PDB by the Joint Centre for Structural Genomics (PDB code 1VR6). This protein was expressed with a non-cleavable N-terminal His-tag. A comparison of this structure with that of the Tyr-bound structure determined in these studies, indicates that this extension at the N terminus has resulted in the elongation of both $\beta1^{r}$ and $\beta2^{r}$ with His (-1 and -3) contributing to interactions with the opposing regulatory domain, inadvertently locking the protein in the closed Tyr-bound conformation (Figure 4.8).

4.4 Small Angle X-ray Scattering

SAXS experiments were employed to assess the solution structure of *Tma*-DAH7PS in comparison to the crystal structures, and to assess the conformational changes upon Tyr binding. Scattering data were collected in the presence and absence of Tyr (Figure 4.9).

For the scattering profile in the absence of Tyr, the radius of gyration, (R_{g}), as determined by Guinier analysis, was 34.1 $\pm$ 0.1 Å (Figure 4.10a). The pair distance distribution function $P(r)$ was calculated using the indirect Fourier transform method. The R_{g} from the $P(r)$ analysis was calculated to be 33.4 $\pm$ 0.1 Å and the maximum dimension of the scattering particle (D_{max}) was approximately 97 Å (Figure 4.10c), which is in close agreement with the open Tyr-free crystal structure (PDB code 1RZM, D_{max} is approximately 99 Å). These scattering profile data for wild-type *Tma*DAH7PS are consistent with the tetrameric form of the protein observed in the crystal structure and predicted for the protein in solution.[39,64]

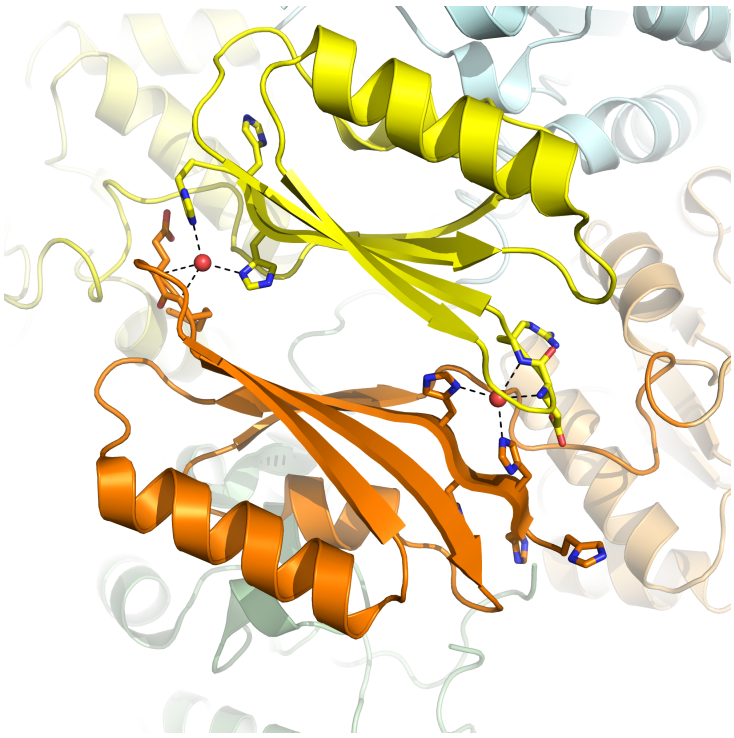
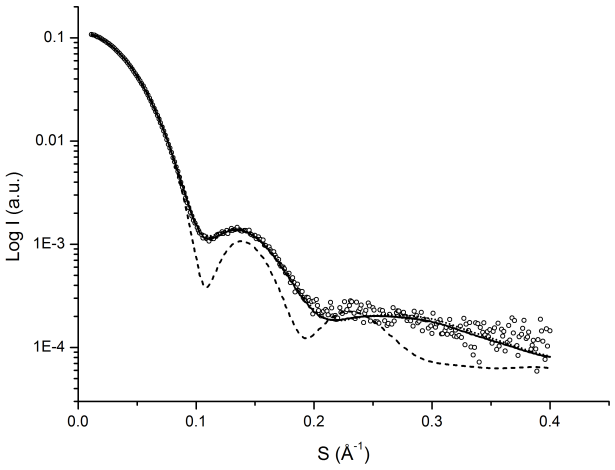


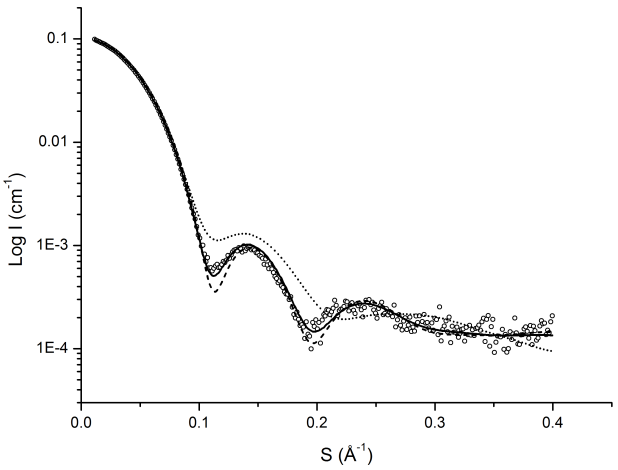
Figure 4.8: Regulatory domain associations in His-tagged *Tma*DAH7PS (PDB code 1VR6). Domains from diagonally opposite regulatory domains in the tetramer are coloured yellow and orange. Water molecules are shown as red spheres.

In the presence of Tyr the R_g, as determined by Guinier analysis, was 32.8 ± 0.1 Å (Figure 4.10b). The R_g from the $P(r)$ analysis was calculated to be 32.0 ± 0.1 Å and the D_{max} was approximately 95 Å (Figure 4.10d). Again, this is in close agreement with the Tyr-bound, closed crystal structure (approximately 92 Å). Taken together, the solution SAXS data suggest in the presence of Tyr *Tma*DAH7PS has a more compact structure than it adopts in the absence of Tyr, a result that is entirely consistent with the crystal structures.

Direct comparisons were made of the open and closed crystal structures to the solution SAXS data using CRYSOL[134] and OLIGOMER.[135] OLIGOMER fits an experimental scattering curve from a multicomponent mixture, in this case the open and closed forms of *Tma*DAH7PS, to find the fraction of each component present in the sample. In the absence of Tyr,



(a)



(b)

Figure 4.9: SAXS measurements of *Tma*DAH7PS, with and without Tyr. Theoretical scattering profiles were generated by using the crystallographic coordinates of both the Tyr open structure (PDB code 1RZM, short dots), and the Tyr-bound, closed structure (PDB code 3PG9, short dashes) using CRYSOL. These theoretical plots were compared to the experimentally determined *Tma*DAH7PS data (○) in both the (a) absence and (b) presence of Tyr. The experimental SAXS data were fitted to a distribution of both the open and closed forms using OLIGOMER (solid line). χ^2_ν values for the fit to the calculated CRYSOL data for the open (1RZM) and closed (3PG9) conformations and values for the OLIGOMER fit for the data in (a) are 0.48, 11.4 and 0.38 respectively, and the values for (b) are 11.0, 1.31, and 0.71.

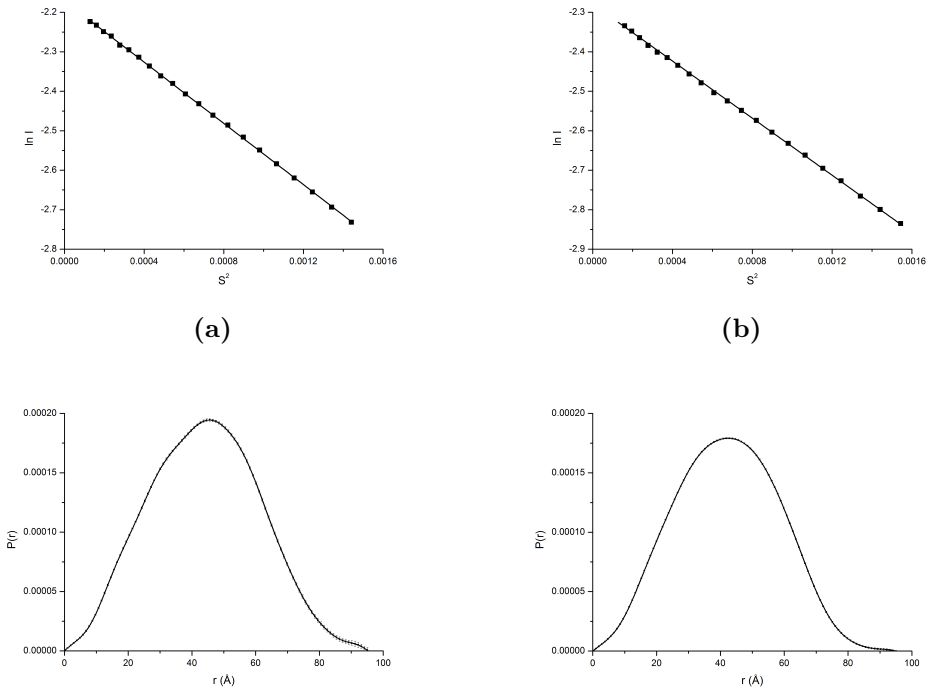


Figure 4.10: Guinier plots and $P(r)$ function calculated for *Tma*DAH7PS in the presence and absence of Tyr. Guinier plots were linear for $s \cdot R_g < 1.3$ for data collected in the (a) absence and (b) presence (b) of Tyr. $P(r)$ function calculated using GNOM[125] from *Tma*DAH7PS SAXS data (c) without and (d) with Tyr. Error bars indicate uncertainty in $P(r)$, propagated from the $I(s)$ vs s profile.

the scattering profile largely fits that calculated for the open *Tma*DAH7PS crystal structure (PDB code 1RZM) when analysed with CRYSOL (reduced chi-squared (χ^2_ν) of 0.48, Figure 4.9a). However, analysis with OLIGOMER predicts that approximately ninety percent of the protein adopts the open form, and approximately ten percent adopts the closed form (χ^2_ν of 0.38). The OLIGOMER fit to this mixture represents a statistically better fit to the data than a fit to the CRYSOL profile generated from the open structure alone ($P_F(F;\nu 1,\nu 2)$<0.05). This analysis is consistent with both open and closed forms of the protein being accessible to the unliganded protein. It is likely that the protein exists as a conformational ensemble, including intermediate conformations between the open and closed conformational extremes observed

in the X-ray crystal structures.

There is a distinct difference in the scattering profile of *Tma*DAH7PS when Tyr is present (Figure 4.9). Here, the scattering data fit more closely to the theoretical profile calculated for the closed, Tyr-bound structure when analysed with CRYSOL (χ^2_ν of 1.3). However, a better fit is obtained once again when the data are analysed with OLIGOMER, which predicts that approximately eighty percent of the protein is present as the closed, inactive form and approximately twenty percent is in the open and active form (χ^2_ν of 0.70). When compared with the CRYSOL fit to the closed Tyr-bound structure, again the fit of the experimental data to the OLIGOMER prediction is statistically better ($P_F(F;\nu 1,\nu 2)<0.05$). The mixture of open (approximately twenty percent) and closed (approximately eighty percent) forms suggested by the scattering of *Tma*DAH7PS in the presence of Tyr, is consistent with the Tyr inhibition results (Figure 2.4), showing that the enzyme displays partial inhibition with respect to Tyr.

These solution SAXS data clearly corroborate the two distinct conformations observed for wild-type *Tma*DAH7PS in the crystal structures. Moreover, the data indicates that Tyr binding favours the adoption of the closed conformation for *Tma*DAH7PS; that is, it latches the regulatory domain into a position that occludes the active site of the adjacent subunit.

4.5 Discussion

The ability to regulate enzymic function is essential for controlling metabolism. As advances towards understanding the finer detail of molecular structure are occurring, alongside is the discovery of the mechanisms of enzyme catalysis. Numerous mechanisms have been reported that confer regulation via ligands binding remotely to active sites.[136] This variety is well demonstrated by the enzyme catalysing the first reaction on the way to aromatic compounds, namely DAH7PS. These mechanisms differ significantly from the subtle synergistic inhibition observed in *Mtu*DAH7PS,[66,106] where a combination of

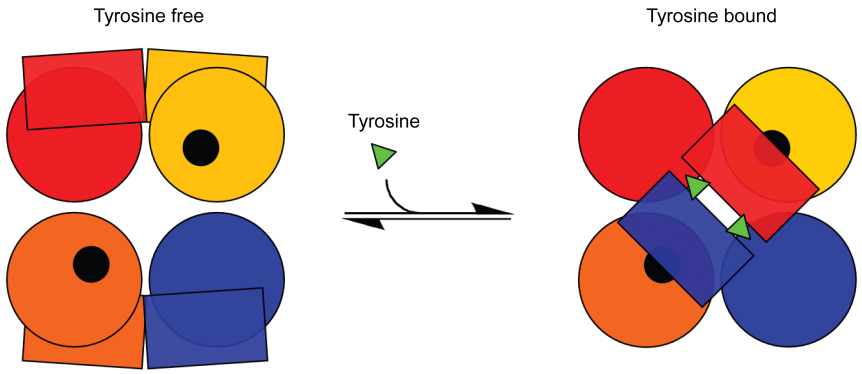


Figure 4.11: Schematic representation of regulatory domain shift of *Tma*DAH7PS on Tyr binding. Tyr (green triangle) binding results in a major structural reorganisation so that the regulatory domains of diagonally opposite molecules occlude the active site (black shading) of the neighbouring barrel. The barrel of each monomer is represented by a circle and the regulatory domain as a rectangle.

the binding of two aromatic amino acids results in minimal structural changes, to the substantial domain reorganisation exhibited by *Tma*DAH7PS in these studies.

For *Tma*DAH7PS, the $(\beta/\alpha)_8$ barrel housing the core catalytic machinery is fused at the N terminus to a distinctive domain with a $\beta\alpha\beta\beta\alpha\beta$ fold. This domain fold is remarkably similar to the ACT domain, which has been acquired, through gene fusion, by a number of other enzymes involved in amino acid biosynthesis for the purpose of conferring regulation.[130,137]

The structure of *Tma*DAH7PS, determined in the presence of Tyr, reveals pairs of Tyr molecules on each side of the tetramer, nestled in binding sites formed by pairs of regulatory domains that extend from different monomers. Sequence alignments suggest that the regulatory domain of *Tma*-DAH7PS is found in a number of other DAH7PS enzymes, indicating that fusion to this domain is a general mechanism for the acquisition of allosteric control by this enzyme.

The molecular details of the allosteric mechanism are evident from structural studies of *Tma*DAH7PS in the presence and absence of Tyr in both crystal and solution forms. Inhibition mediated by Tyr is shown to be associated with a considerable reorganisation of the structure of the enzyme. The enzyme can adopt both an open form, where the regulatory domains extending from each monomer of the tetrameric protein show no association, or a closed form, in which the C-terminal end of the barrel housing the active site is capped, and substrate access is restricted. A dynamic equilibrium between the open and closed forms exists. Tyr perturbs the position of this equilibrium, effectively latching the swinging regulatory gate into the closed position (Figure 4.11). In the unliganded open form, transient associations result in the potential for the regulatory domains to interact. These interactions may help stabilise quaternary structure and provide a clue to why $^{\text{trunc}}$*Tma*DAH7PS is dimeric in crystalline form.

Although the gross domain reorganisation observed on Tyr binding for *Tma*DAH7PS shown in this study appears to be far more dramatic than the more subtle movements associated with allosteric control of other proteins, there are some parallels with other systems.[138,139] As more structural studies reveal the details of both liganded and unliganded forms of proteins, it seems likely that similar strategies for allosteric control will be uncovered. What is becoming clear is: (1) that ACT domains typically combine; (2) that ligand-binding sites are frequently provided at the interfaces between two or more ACT domains; and (3) that associations between these domains are integral to the mechanisms of allosteric control. For example, binding of histidine to ATP phosphoribosyl transferase, the enzyme that catalyses first committed step of histidine (His) biosynthesis, results in perturbation of the equilibrium position between active dimeric and inactive hexameric forms. This change is mediated via interactions between ACT domains.[138] Likewise, tetramerisation of the P-protein (with chorismate mutase and prephenate dehydratase activities) involves regulatory ACT domains and is associated with enzyme inhibition.[139] It has been suggested that alteration in quaternary interactions may prove to be a general mechanism for allosteric control for

proteins that bear this regulatory extension.[130]

The quaternary structure of *Tma*DAH7PS is essential for allosteric control. There are significant interactions across the tetramer interface when the protein adopts the closed conformation, not only between adjacent monomers but also between opposite monomers involving both the catalytic and regulatory domains. These studies provide insight into the interactions between quaternary structure and regulatory extensions that deliver regulated catalysis of *Tma*DAH7PS. Regulation is achieved by the abrogation of substrate entry to the active site and is mediated by a physical gating mechanism controlled by Tyr binding.

In conclusion, these studies illuminate an intricate interplay between quaternary structure and regulatory extensions that delivers the regulated catalysis of this enzyme. Regulation is achieved by a remarkable ligand-controlled gating mechanism governing substrate access to the enzyme active site.

Chapter 5

Investigating the Residues in *Thermotoga maritima* DAH7PS Important for Ligand Binding and Regulation

5.1 Introduction

The structural and functional characterisation of a truncated variant, lacking the N-terminal domain (Chapter 3), and the Tyr-bound structure of *Tma*-DAH7PS (Chapter 4) have shown that this extra structural motif decorating the basic $(\beta/\alpha)_8$ barrel of *Tma*DAH7PS is essential for allosteric regulation. A significant conformational change of the tetrameric protein occurs on ligand binding; whereupon Tyr is bound in a cleft formed between two opposing N-terminal ACT domains. This ligand-binding site physically obstructs substrate entry to the active site, impeding the catalytic activity of the enzyme.

Previous studies of the type Iα DAH7PS subfamily have proposed that

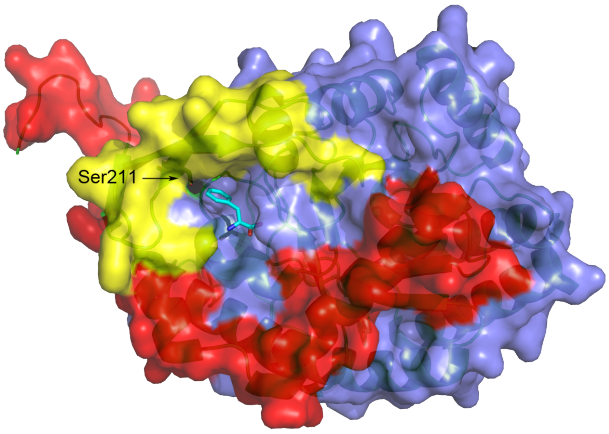


Figure 5.1: Surface representation of *Eco*DAH7PS (PDB code 1KFL) showing how the N-terminal (red) and internal extension (yellow) together create a binding cavity for the inhibitor, Phe (cyan, sticks). The specificity determining Ser residue is shown as green sticks.

specific feedback sensitivity can be predicted by the inspection of primary sequence.[33] This DAH7PS subfamily has two extensions to the core TIM-barrel structure, which together create a binding cavity for the inhibitor molecule (Figure 5.1).

It has been proposed that one key residue within this binding cavity dictates the sensitivity and ultimately defines the specificity of the enzyme to inhibition by Phe and Tyr.[33] Specifically, for the Phe-regulated isozyme of *Sce*DAH7PS, and all Phe sensitive Iα DAH7PSs, a hydroxyl group is provided by a conserved Ser residue found at the bottom of the inhibitor-binding cavity. The smaller conserved Gly residue found in this same position for the Tyr-regulated isozyme increases the size of the cavity and accommodates the steric bulk of the Tyr ligand's hydroxyl group. The single mutation of Gly-to-Ser in Tyr-regulated DAH7PSs leads to a complete switch in regulatory ligand specificity without affecting kinetics.[33]

In the type Iβ *Tma*DAH7PS enzyme, the assembled regulatory domains also create an inhibitor-binding cavity, with contributions by residues from the regulatory domains of two different monomers. To investigate the essential

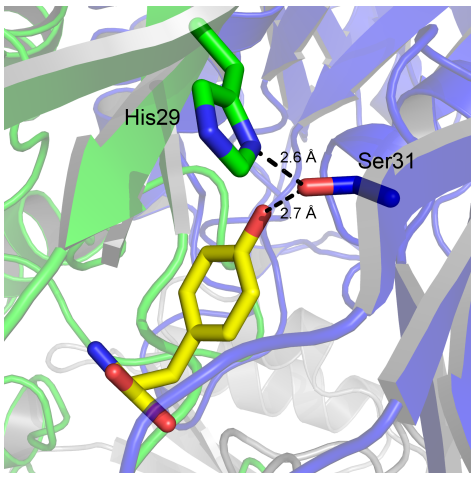


Figure 5.2: Hydrogen-bonding interactions (black dashes) of residues His29 (green sticks) and Ser31 (blue sticks) from opposing regulatory domains with Tyr (yellow sticks).

binding interactions of Tyr and Phe with *Tma*DAH7PS, variants were designed by analysis of the ligand-bound crystal structure of *Tma*DAH7PS, targeting residues with side-chain ligand interactions to the bound Tyr.

One of the notable hydrogen-bonding interactions with Tyr in the Tyr-bound structure of *Tma*DAH7PS is between the side chain of Ser31 and the hydroxyl group of the Tyr ligand (Figure 5.2). His29, contributed from the opposing regulatory domain, hydrogen bonds to the hydroxyl group of Ser31 and is also in close proximity to the hydroxyl group of the bound Tyr. Although possibly not directly involved in binding to the inhibitor, His29 may be important for securing the correct orientation of Ser31 or the ligand.

Could a single mutation be enough to change inhibitor specificity in *Tma*DAH7PS? Furthermore, is the absolute positioning of His29 and Ser31 important? This chapter describes the investigation into the importance of key interactions between the inhibitor molecule and the residues of the enzyme that may impart inhibitor specificity.

5.2 Choice of Mutations

The mutational changes were chosen based on their interaction with Tyr and by analogy with the studies of the Iα DAH7PS subfamily, in which ligand specificity appears to be determined by a single residue.[33] Unlike the Tyr-sensitive Iα DAH7PS enzymes, which have a Gly residue in the inhibitor-binding cavity, *Tma*DAH7PS has a Ser residue, similar to the Phe sensitive enzymes. This Ser residue appears to be important because it forms a hydrogen bond to the hydroxyl group of Tyr. Ser31 was mutated to Gly (creating *Tma*S31G), to remove the specific interaction with Tyr, and to also eliminate the hydrogen bond to His29 on the opposing subunit.

Although His29 has no direct interaction with Tyr, it is in close proximity to the ligand and forms a hydrogen bond with Ser31. Mutation of this His to alanine (Ala) (creating *Tma*H29A) was designed to remove any hydrogen bonding capabilities of the residue's side chain, and replace it with a small non-intrusive side chain. Analysis of this variant protein may provide insights into the importance of the interdomain contacts between His29 and Ser31, and the secondary role His29 may play in assisting with correct placement of Ser31. Furthermore, to probe the importance of possible hydrogen-bonding contacts between the two domains, mediated by His29 and Ser31, His29 and Ser31 were interchanged and the resulting variant (*Tma*H29S/S31H) analysed for effect of this substitution on regulation/inhibition.

5.3 Mutagenesis, Expression and Purification

A QuikChange® Lightning Site-Directed Mutagenesis Kit was used to introduce site-specific variants in the *Tma*DAH7PS pT7-7 plasmid. Extension of primers, containing the desired mutations, generated mutated plasmids. The PCR product was subjected to DpnI digestion, agarose gel analysis

and transformation into *E. coli* TOP10 cells. Small cultures of individual colonies were grown and plasmids extracted. DNA sequencing confirmed the identity of colonies containing the desired mutations. Plasmids with the correct sequence were transformed into the protein expression cells *E. coli* BL21 Star™(DE3). Cultures were grown, protein expression induced, and over-expressed protein purified using the same method as was developed for the purification of wild-type *Tma*DAH7PS (Chapter 2, Sections 2.2 and 2.3).

5.4 Confirmation of Molecular Weight

The molecular weights of the variants proteins were determined by electrospray ionisation mass spectrometry, and are in close agreement with the molecular weights calculated from the amino acid sequence (Table 5.1).

Table 5.1: Expected and experimentally determined molecular weights of *Tma*DAH7PS variant proteins

*Tma*DAH7PS	Expected (Da)	Measured (Da)
H29A	37 312	37 311
S31G	37 348	37 348
H29S/S31H	37 378	37 378

5.5 Kinetic Parameters

All three variants were found to be catalytically active and displayed similar kinetic parameters to those of wild-type *Tma*DAH7PS. The steady-state kinetic constants were determined for each of the variants using the same assay method as was employed for wild-type *Tma*DAH7PS (Chapter 2, Section 2.5). In comparison to wild-type *Tma*DAH7PS, the variants all display higher $K_{\mathrm{m}}^{\mathrm{PEP}}$ and $K_{\mathrm{m}}^{\mathrm{E4P}}$ values despite the mutations being distant to the active site

Table 5.2: Kinetic parameters for *Tma*DAH7PS variants

*Tma*DAH7PS	K_m^{PEP} (µM)	K_m^{E4P} (µM)	k_{cat} (s^{-1})
WT	4.85 ± 0.04	13 ± 1	11.7 ± 0.2
H29A	12.8 ± 0.4	30 ± 1	17.5 ± 0.3
S31G	21 ± 2	33 ± 2	14.3 ± 0.2
H29S/S31H	39 ± 1	28 ± 1	14.1 ± 0.1

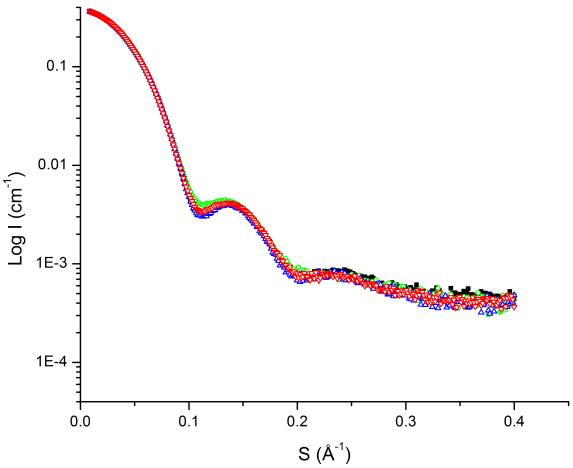
(Table 5.2). Interestingly, the variants also show slightly higher k_{cat} values than wild-type *Tma*DAH7PS. However, for thermophilic DAH7PS enzymes, some inconsistencies in k_{cat} values recorded at elevated temperatures have been previously noted (Schofield, unpublished data).

5.6 Small Angle X-ray Scattering

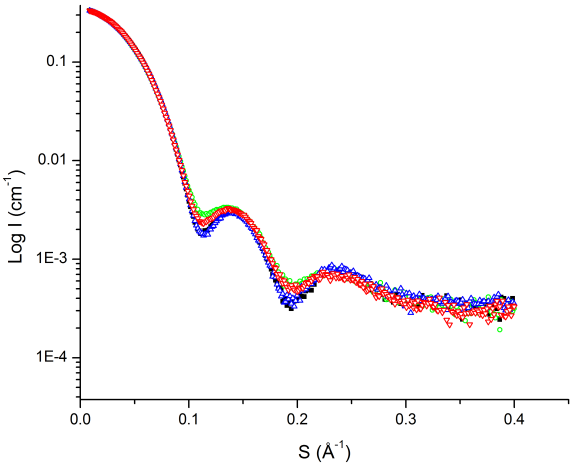
SAXS was used to investigate whether the conformational change observed on inhibitor binding for wild-type *Tma*DAH7PS is also apparent for *Tma*H29A, *Tma*S31G and *Tma*H29S/S31H. Data were collected in the absence and presence of Tyr or Phe, and the scattering profiles collected were compared to the response of wild-type *Tma*DAH7PS.

The scattering profiles for all proteins were very similar in the absence of Tyr or Phe, indicating that all variants adopt (in solution) the same overall fold and tetrameric quaternary structure as wild-type *Tma*DAH7PS (Figure 5.3a). There appears to be a slight variation in the scattering profile of *Tma*S31G, which perhaps is an indication of the open/closed equilibrium being even more in favour of the open form.

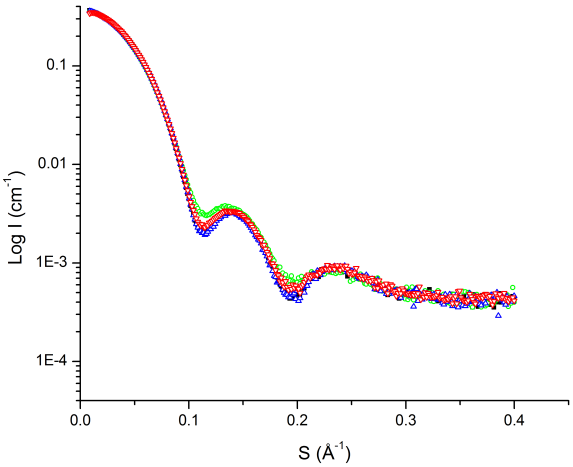
On the addition of Tyr or Phe, significant changes to the scattering profiles were observed, in particular the scattering around $0.1\,Å^{-1}$ and $0.2\,Å^{-1}$ (Figure 5.3). This change is in line with the profile obtained when Tyr is added to wild-type *Tma*DAH7PS, and suggests that in the presence of either Tyr or Phe, all variants undergo a structural change to acquire a more globular



(a)



(b)



(c)

Figure 5.3: Comparison of SAXS measurements from *Tma*H29A (△), *Tma*S31G (○) and *Tma*H29S/S31H (▽), compared to those collected from wild-type *Tma*DAH7PS (■) in the (a) absence of ligand, (b) presence of Tyr or (c) presence of Phe.

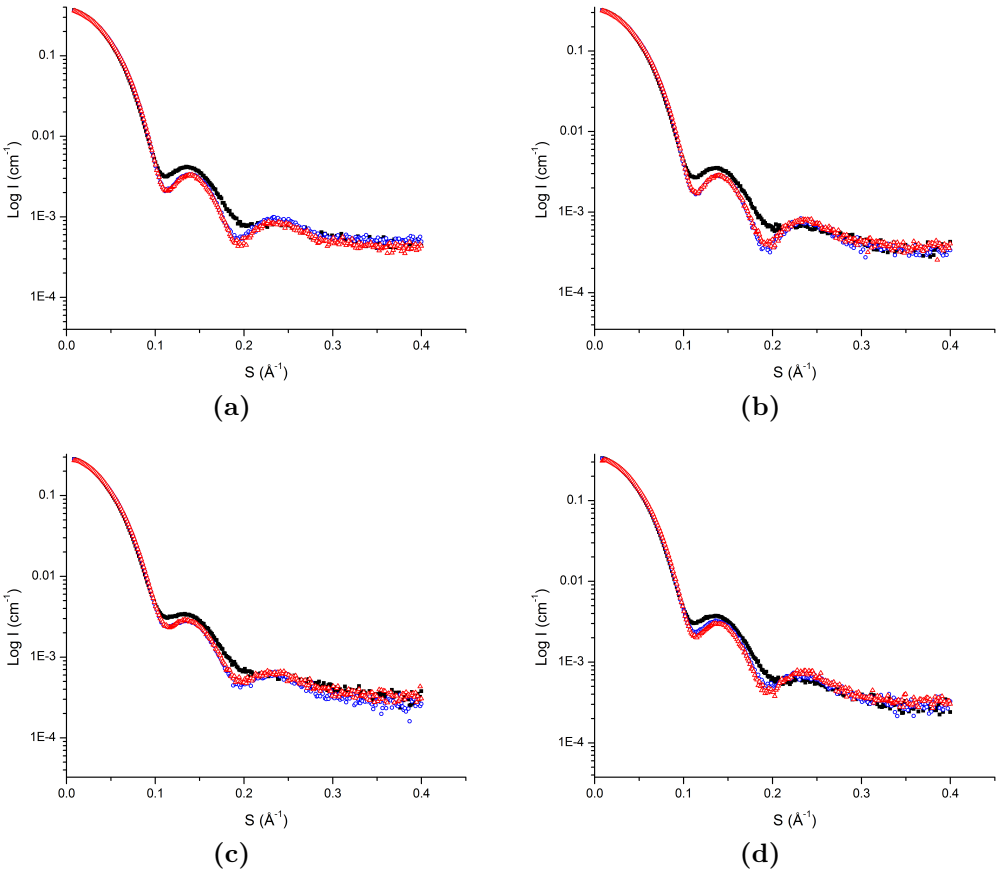


Figure 5.4: SAXS measurements of (a) wild-type *Tma*DAH7PS, (b) *Tma*H29A, (c) *Tma*S31G, and (d) *Tma*H29S/S31H comparing SAXS measurements in the absence of inhibitor (■), presence of Tyr (○) and presence of Phe (△).

closed conformation similar to that adopted by wild-type *Tma*DAH7PS.

The changes to the scattering profile for *Tma*H29A in the presence of Tyr or Phe (Figure 5.4b), mirror closely the changes observed for wild type, suggesting the presence of these inhibitors at the concentrations employed elicits a similar degree of conformational change in both proteins. However, for *Tma*S31G in the presence of Phe and Tyr, the changes in the scattering profile are similar but are not as pronounced as the *Tma*H29A construct. Although the overall shape adopted by *Tma*S31G appears to be more globular in the presence of Tyr or Phe, close analysis of the SAXS measurements reveals a flatter profile (Figure 5.4c) than for wild-type protein and *Tma*H29A,

suggesting this variant adopts a more elongated structure.[140] This result may indicate some impairment to the regulatory domain interaction in this particular variant.

A similar response to that observed for *Tma*S31G is also observed for *Tma*H29S/S31H (Figure 5.4d). As a dynamic equilibrium is likely to exist between the open and closed forms of *Tma*DAH7PS, and as SAXS data are derived from an average structure in solution, removing the Ser31 hydroxyl may have disturbed this balance, pushing the protein more towards the open conformation, resulting in the change observed in the scattering profile. Generally, the addition of Tyr and Phe to the variant enzymes generate the same conformational response in solution, however for *Tma*H29S/S31H a slight increase in change for Phe over Tyr is apparent.[i]

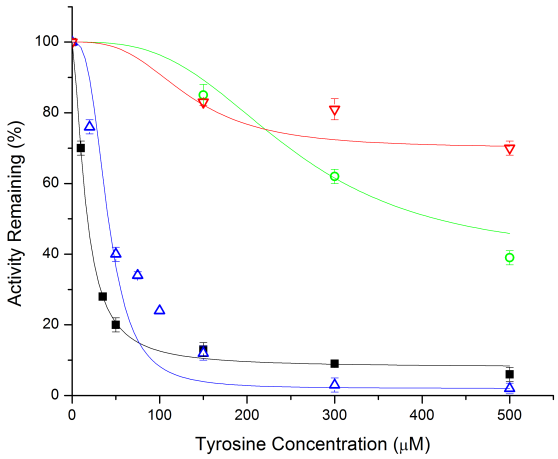
5.7 Response to Inhibitors

Table 5.3: Estimated IC50 values for wild-type and variant *Tma*DAH7PS

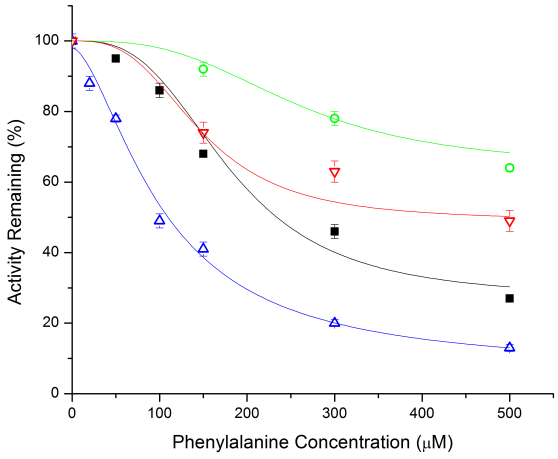
*Tma*DAH7PS	Estimated IC50 (µM)	
	Tyr	Phe
WT	20	180
S31G	250	260
H29A	40	100
H29S/S31H	130	150

The degree of sensitivity towards Tyr or Phe for each variant was tested using increasing concentrations of each inhibitor (Figure 5.5). The analysis of this inhibition is quite complex and several tools were employed to help gain insight into the variable responses of these enzymes to inhibitors. The half-maximal inhibition concentrations (IC50) were calculated using the Logistic Fit (Origin Pro 8.1) and these values were directly compared to the response

[i] The $P(r)$ functions and Guinier plots calculated from the variant protein SAXS data are located in the appendix.



(a)



(b)

Figure 5.5: Comparison of *Tma*H29A (△), *Tma*S31G (○), *Tma*H29S/S31H (▽) and wild-type *Tma*DAH7PS (■) response to varying concentrations of (a) Tyr and (b) Phe. Error bars indicate standard deviation. Sigmoidal curves were generated and estimated IC50 values were determined by fitting the data to the Logistic Model using OriginPro 8.1.

of the wild-type protein (Table 5.3). These IC50 values hide the fact that the total inhibition is greater for some variants at higher ligand concentrations, therefore for greater clarity the levels of inhibition at 300 µM of inhibitor were also directly compared (Figure 5.6).

Tyr does not inhibit *Tma*S31G to the same extent as it inhibits wild-type *Tma*DAH7PS. Similarly, *Tma*S31G is slightly less sensitive to Phe than the wild-type protein. Notably, the selectivity for Tyr over Phe is all but abolished, with *Tma*S31G showing similar levels of inhibition with either aromatic amino acid. These results reveal the importance of Ser31 for both sensitivity toward Tyr and inhibitor specificity.

Of the three variants, *Tma*H29A behaved most similarly to the wild-type enzyme with a comparable response in the presence of Tyr. However, this protein shows a significantly increased response to Phe. In fact, Phe is a more potent inhibitor for *Tma*H29A than for any of the other proteins. Intriguingly, this indicates that while sensitivity is not lost, selectivity for Tyr over Phe was compromised.

*Tma*H29S/S31H, shares some similarities to *Tma*S31G, in that inhibition by Tyr is significantly compromised. Interestingly, this protein was found to

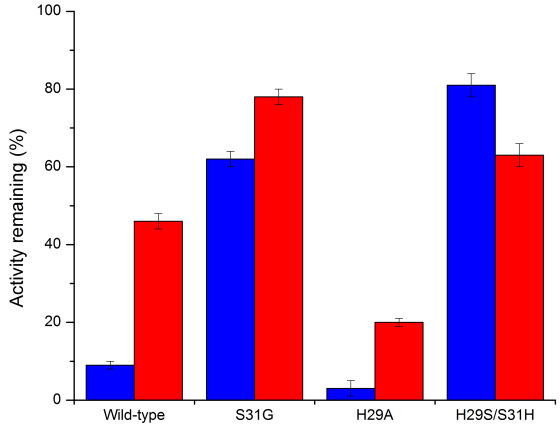


Figure 5.6: Wild-type and variant *Tma*DAH7PS response to 300 µM Tyr (blue) and Phe (red).

respond to Tyr and Phe with similar sensitivity.

5.8 Discussion

Some DAH7PS enzymes, including *Mtu*DAH7PS,[106] *Eco*DAH7PS[27] and *Sce*-DAH7PS[33] have a pre-formed ligand-binding cavity and undergo very subtle structural changes on ligand binding. However, the N-terminal regulatory domain of *Tma*DAH7PS, a type Iβ enzyme, undergoes a significant change in conformation to create a ligand-binding site. Residues from two different protein chains contribute to ligand binding, and direct inter-subunit contacts are formed. To elucidate the role of different residues in determining the sensitivity and selectivity of the allosteric response in *Tma*DAH7PS, mutations were created to disrupt key interactions between the protein and bound inhibitor.

These studies have shown that Ser31 plays a significant role in inhibitor binding for *Tma*DAH7PS. This is not the first time a Ser residue has been shown to be important in the allosteric regulation of a DAH7PS enzyme.[141] Although Ser is not charged at physiological pH, the primary OH group is hydrophilic and forms hydrogen bonds. Due to its ability to act as either a donor, acceptor group, or both, it is found in diverse electrostatic environments.[142] Hydroxyl groups have a degree of rotational freedom that allows them to align themselves to minimise the energy of their hydrogen-bonding interactions. The observed rotamer conformation indicates the predominance of either the strongest hydrogen-bond formation or the arrangement providing the lowest steric effect.

The versatility of Ser is evident in the Tyr-bound crystal structure of *Tma*DAH7PS where the alcohol containing side chain of Ser31 forms a hydrogen bond with the hydroxyl moiety of the bound Tyr ligand. Furthermore, Ser31 also forms a hydrogen bond with His29 from the opposing regulatory domain, which may ensure the correct positioning of Ser31 to bind Tyr and the correct construction of the binding pocket. Exchange of the

neutral-polar side chain of Ser31 for the simple hydrogen of Gly, eliminates the hydrogen-bonding ability of the side chain and severely compromises the inhibition of *Tma*DAH7PS by Tyr. Interestingly, this mutation also slightly decreases the inhibition of *Tma*S31G by Phe in comparison to the wild-type enzyme, but Tyr is still more potent than Phe for inhibition of this modified enzyme. This observation suggests that the hydrogen-bonding capability of the hydroxyl side chain of Ser31, with the hydroxyl group of the Tyr ligand and with His29, plays an important role for selection between Tyr or Phe. Preliminary molecular dynamics results suggest that the replacement of the hydroxymethyl side chain of Ser with a hydrogen provides sufficient space for a water molecule to be recruited to hydrogen bond with the Tyr hydroxyl group (Jiao, personal communication). This may provide a plausible explanation for the observation that domain closure still takes place, albeit poorly, and that *Tma*S31G demonstrates slightly greater sensitivity to Tyr than Phe. The overall loss of responsiveness towards Tyr and Phe indicates that Ser31 is also important for the inhibitor sensitivity. Furthermore, the SAXS scattering profile for *Tma*S31G indicates that a conformational change similar to that observed for wild-type occurs on inhibitor binding. However, the change is more subtle and substantiates the loss of sensitivity towards both Tyr and Phe.

The ability of *Tma*H29A to hydrogen bond to Tyr is presumed to be preserved in this variant by the presence of Ser31. Although the removal of the imidazole functionality does not appear to affect the enzyme's responsiveness to Tyr in comparison to wild-type, a dramatic increase in response to Phe is observed. In fact, this enzyme shows the highest sensitivity towards Phe of any of the proteins. This is consistent with the importance of the hydrogen bond between Ser31 and His29 for inhibition of the enzyme by Tyr, and suggests that His29 plays a role in determining selectivity towards Tyr. The increase in sensitivity to Phe is somewhat puzzling. If the hydrogen bond to Ser31 was important to the hydrogen-bonding network, as suggested by the S31G variant, one would expect *Tma*H29A to be similarly impaired. Crystal structures of these variants with bound ligand may help clarify reasons for

this apparent anomaly.

It is well established that aromatic residues undergo non-covalent energetically favourable interactions between side chains of the amino acids Phe, Tyr, Trp and, although complicated by the positive charge, of His as well.[143] Aromatic-aromatic interactions are prevalent in many areas of chemistry.[144] Examples include the aggregation of porphyrins,[145] conformation of phenylactetylene macrocycles,[146] and the strength of kevlar can be in part be attributed to aromatic-aromatic interactions.[147] In biological chemistry, aromatic-aromatic interactions are crucially involved in protein DNA complexes where the interactions between aromatic side chains and base pairs are actually visible in X-ray structures.[148,149]

Some studies have suggested that phenyl-rings align themselves in a perpendicular arrangement, whereas others report that pairs or dimers of aromatic acids prefer to align their respective aromatic rings in an off-centred parallel arrangement consistent with those found for the benzene dimer (Figure 5.7a).[144] Either way, it is obvious that dimers of these residue side chains have many possible stabilising interactions and at distances larger than the average van der Waals radii.[144] It has been suggested that sixty percent of aromatic side chains in proteins are found to be involved in aromatic-aromatic interactions and up to eighty percent of these energetically favourable albeit weak aromatic-aromatic interactions stabilise protein quaternary structure.[150]

In the Tyr-bound crystal structure of *Tma*DAH7PS, Tyr and His29 appear to adopt a T-shaped π-stacking configuration (Figure 5.7b) and the loss of this interaction may explain the small reduction in sensitivity towards Tyr. However, the aromatic-His interactions with Phe or with Tyr have been reported as being inequivalent, with the interaction of Tyr and His less stabilising than the Phe-His interaction.[151] From this it would be expected that disturbance of the His-Phe interaction would have a larger effect than His-Tyr. However, the increase in sensitivity observed in *Tma*H29A towards Phe argues against the role of π-stacking for His29. The significant increase in sensitivity toward Phe may simply be due to the increased opportunity to access the ligand-binding site resulting from removal of the steric bulk of the

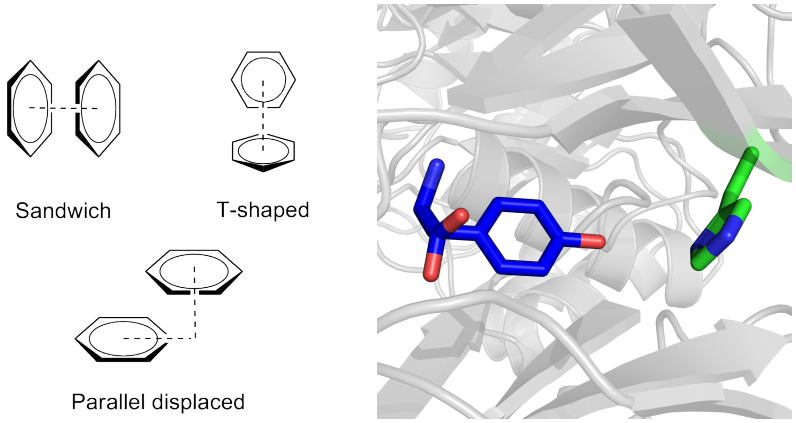


Figure 5.7: (a) Representative conformations of the π interactions observed in the (a) benzene dimer compared with the possible equivalent interaction in (b) wild-type *Tma*DAH7PS

imidazole side chain of His29. Alternatively, replacement of the slightly polar His residue with Ala may create a more hydrophobic environment, favouring binding of Phe.

As hydrogen-bonding interactions will be disturbed by the simultaneous exchange of His29 and Ser31, a large effect on estimated IC50 values may have been expected. Unsurprisingly, this mutational pairing severely compromises selectivity and sensitivity to inhibition by Tyr. Assuming Tyr remains close to its position in wild-type *Tma*DAH7PS, analysis of the possible rotamers accessible in these exchanged positions suggest there is no arrangement that would permit a hydrogen bond between the hydroxyl group of Tyr and Ser29. However, a hydrogen bond could be made between the Tyr hydroxyl and His31 and an extended hydrogen bond between His31 and the main chain of Ser29 (Figure 5.8).

Nevertheless, inhibition results for *Tma*S31G and *Tma*H29S/S31H suggest the hydrogen bond, specifically to the Tyr hydroxyl group, is compromised in line with the similar abilities of Tyr and Phe to inhibit this variant. These results are further reflected by the SAXS scattering profiles obtained for these two variants, in the absence and presence of Tyr and Phe, which show a

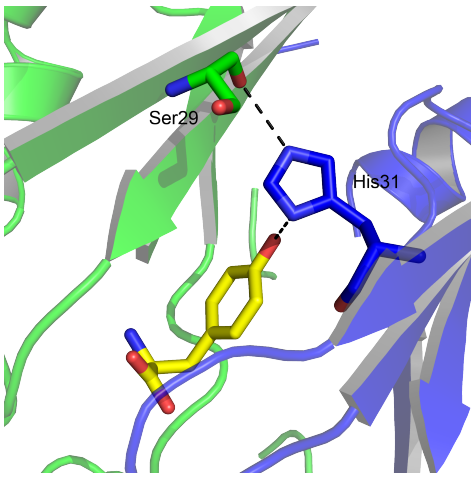


Figure 5.8: Potential hydrogen-bonding interactions of residues His31 (blue sticks) and Ser29 (green sticks) from opposing regulatory domains with Tyr (yellow sticks).

decreased response to inhibitor-induced conformational change.

Although Ser31 is not conserved across the ACT domain of the Iβ DAH7PSs that possess this N-terminal extension, mutational studies of *Tma*-DAH7PS have shown the importance of this residue for selective Tyr binding. Subsequent studies to further probe the importance of this residue could include mutation to an Ala residue. This mutation would eliminate the hydrogen-bonding capabilities of the residue to the inhibitor and preclude water from accessing the cavity and acting as a surrogate hydrogen-bonding partner, therefore allowing a more complete assessment of the importance of this Ser residue. Phe inhibition appears to be more complicated with a number of less obvious contributing factors that are yet to be uncovered. A crystal structure of *Tma*DAH7PS with Phe bound should be sought to help with this analysis.

Chapter 6

Fusion of the Regulatory (ACT) Domain from a Regulated DAH7PS to an Unregulated DAH7PS Barrel

6.1 Introduction

Protein remodelling in order to obtain proteins and enzymes with desirable properties and activities is a current focus of biotechnology.[152] The manipulation of existing protein scaffolds to impart ideal properties can be achieved by combinatorial methods such as molecular breeding,[153] DNA shuffling,[154] or the exchange of elements between two or more existing enzymes to create hybrid enzymes. The elements that can be swapped may be single residues, secondary elements, subdomains, domains or entire proteins.[155]

ACT domains are found in a large number of proteins. A majority of ACT-domain containing proteins appear to interact with amino acids and are involved in some aspect of regulation of amino acid metabolism.[132,156,157]

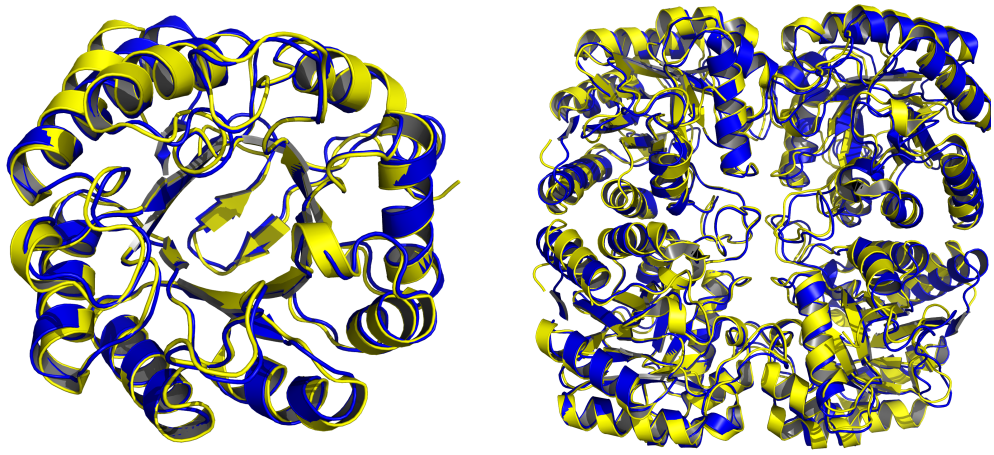
Increasing diversity in tertiary and quaternary architecture and ligand-binding interactions are being uncovered for the ACT domain as more structures are determined. It has been suggested that the recruitment of ACT domains for the purpose of allosteric regulation can be explained by its ability to create symmetric or asymmetric assemblies, resulting in a highly diverse array of binding sites and mechanisms of control.[131]

Most ACT domains are found within Bacteria and Archaea, which indicates that appearance of this domain occurred early in evolutionary history, however there is at least one functional example found in a mammalian protein, rat phenylalanine hydroxylase.[158] There appears to be no distinct consistency of the use of ACT domains within a particular species or within proteins that possess a common activity. It has been suggested that the ACT domain represents a "conserved, evolutionarily mobile module" that when fused to other proteins, makes them susceptible to regulation by small molecules.[132]

The recruitment of the ACT domain in the Iβ DAH7PS subfamily is clear from inspection of structure and sequence. This makes this subfamily the ideal system in which to explore whether the property of allosteric regulation is an intrinsic property of the ACT domain itself, encoded entirely by this domain, or whether an essential interplay exists between the regulatory ACT domain and DAH7PS catalytic barrel. This chapter will explore the mobility of the ACT domain from DAH7PS by fusing the ACT domain (derived from a regulated DAH7PS, *Tma*DAH7PS) onto an unregulated DAH7PS scaffold (*Pfu*DAH7PS).

6.2 Choice of Fusion Site

At the commencement of these studies, *Pfu*DAH7PS was the only thermophilic, unregulated type Iβ DAH7PS enzyme that had been characterised both functionally and structurally.[38,48] Due to the availability of the gene encoding for *Pfu*DAH7PS, familiarity with the enzyme, and the structural similarity between the barrel elements of *Pfu*DAH7PS and *Tma*DAH7PS (Figure 6.1),



(a) **(b)**

Figure 6.1: (a) *Pfu*DAH7PS (PDB code 1ZCO) monomer (yellow) superimposed with the catalytic barrel of *Tma*DAH7PS (PDB code 1RZM) (blue) (RMSD of 0.539 Å) and (b) tetramer overlay (RMSD of 0.520 Å).

*Pfu*DAH7PS was chosen as the scaffold for fusion to the regulatory ACT domain of *Tma*DAH7PS.

The crystal structures of both *Tma*DAH7PS and *Pfu*DAH7PS show the existence of a β-hairpin at the N-terminal end of the barrel. Two constructs of ^{TmaPfu}DAH7PS were designed to preserve the β-hairpin and allow a sufficiently long linker between the ACT domain and catalytic barrel. Construct 1 incorporated the β-hairpin from *Pfu*DAH7PS, which is slightly shorter than the equivalent region in *Tma*DAH7PS. Whereas, construct 2 derived its linker and β-hairpin from *Tma*DAH7PS, and thus more closely resembled *Tma*DAH7PS. The determination of the ligand-bound crystal structure of *Tma*DAH7PS illustrated the importance of the linker for regulation and initial expression trials revealed construct one was insoluble. Therefore, construct one was not considered further in these studies.

The selected ^{TmaPfu}DAH7PS construct incorporated both the β-hairpin and the linker located between the catalytic and ACT domains of *Tma*-DAH7PS. This linker is thought to enable the conformational rotation of the ACT domain that occurs on ligand binding for *Tma*DAH7PS and allows

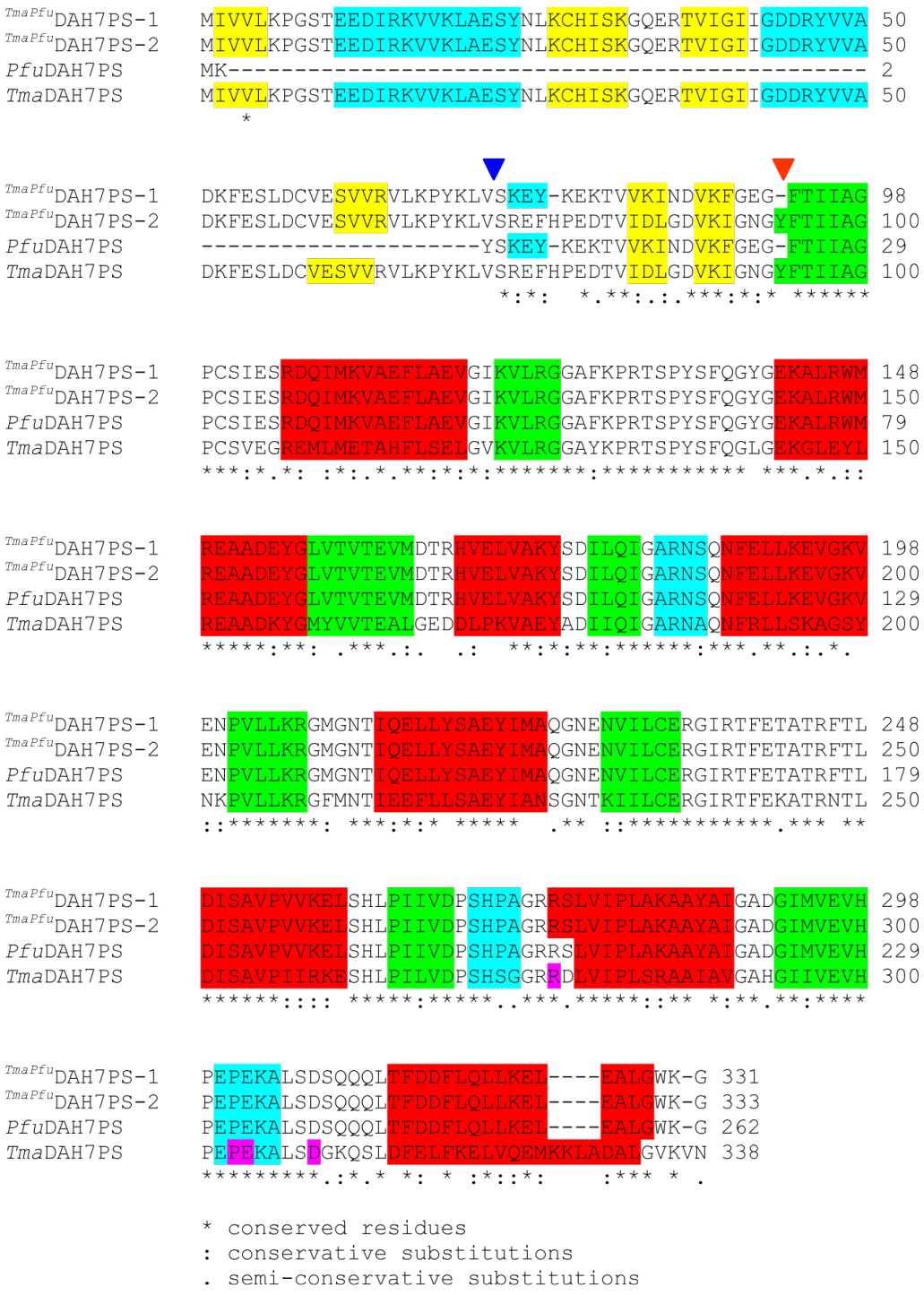


Figure 6.2: Primary sequence of the two constructs of the chimeric protein, *TmaPfu*DAH7PS, comprising of fusion between the catalytic domain of *Pfu*DAH7PS and the ACT domain of *Tma*DAH7PS. The fusion sites for construct 1 (▼) and construct 2 (▼) are indicated. α-helices and β-strands of the barrel are highlighted red and green, respectively. Non-core α-helices are highlighted cyan and non-core β-strands are coloured yellow. The residues of the *Tma*DAH7PS barrel, which are identified as having interactions with the ACT domain on ligand-induced closure, are highlighted pink.

the two domains opportunity to fold correctly to achieve a tertiary structure without interference from one another.[152]

6.2.1 Sequence Conservation

Sequence alignment analysis of the barrel region of all type Iβ DAH7PS enzymes show that the general level of conservation is relatively low with only nineteen conserved residues across this region. Similarly, alignments of the ACT domains found within the type Iβ DAH7PSs show only a few semi-conserved residues. Despite the sequence variation found within the Iβ subfamily of DAH7PS enzymes, the barrel regions of *Pfu*DAH7PS and *Tma*-DAH7PS appear similar in both structure (RMSD of 0.539 Å) and sequence, with 166 conserved residues or 63 percent identity (Figure 6.2). Hence, the choice of *Pfu*DAH7PS and *Tma*DAH7PS as fusion partners potentially parallels the domain recruitment strategy nature uses to evolve enzymes.[155]

6.3 Expression and Purification

A synthetic gene encoding for ^{TmaPfu}DAH7PS was designed and ordered from GeneArt® Gene Synthesis service.[i] The construct was codon optimised for *E. coli*, cloned into the Gateway® entry vector pDONR™ 221 and sequence-verified prior to arrival. An expression clone was generated in an untagged vector, pDEST14, using standard Gateway® technology and this construct was transformed into *E. coli* One Shot® BL21 Star™ (DE3) cells. The expression of ^{TmaPfu}DAH7PS in this cell line produced protein of the appropriate molecular weight as observed by SDS-PAGE. However, the protein predominantly appeared in the cell debris after lysis by sonication (Figure 6.3). The co-expression of different chaperones has been frequently used to help alleviate expression stress and increase the solubility of recombinant

[i] The codon-optimised nucleotide gene sequence of ^{TmaPfu}DAH7PS is located in the Appendix.

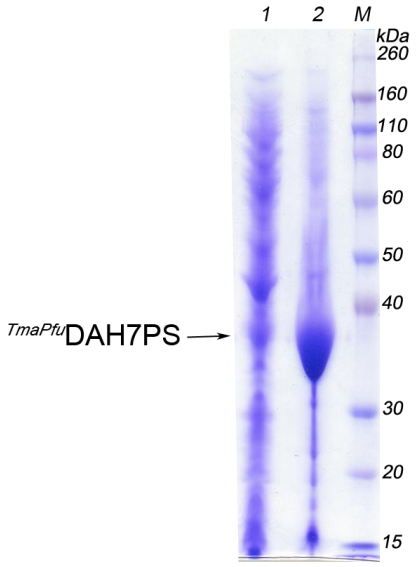


Figure 6.3: SDS-PAGE analysis of protein expression of ^{TmaPfu}DAH7PS from *E. coli* BL21(DE3) cells after lysis in the supernatant (lane 1) and cell debris (lane 2).

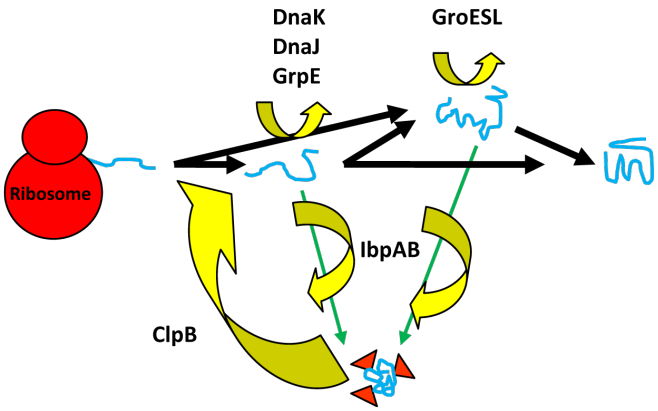


Figure 6.4: Role of molecular chaperones in protein folding and disaggregation. The DnaK system (DnaK, DnaJ, GrpE) in combination with the chaperonins (GroESL) may assist the folding intermediates to reach their native folded structure. During aggregation the small heat shock proteins IbpAB (orange triangles) bind the misfolded proteins and ClpB triggers the disaggregation process. Diagram adapted from deMarco.[159]

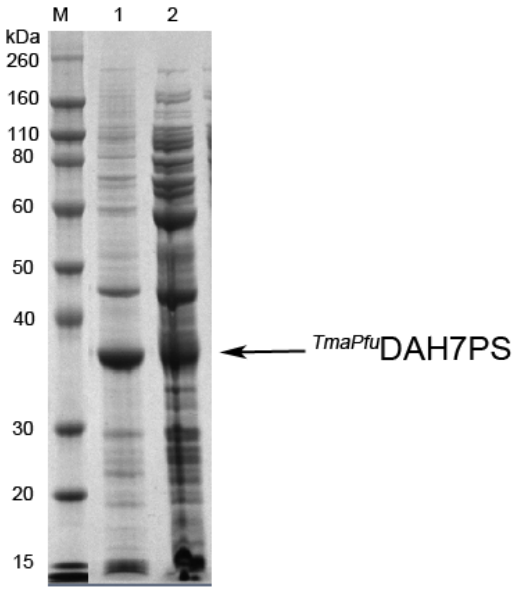


Figure 6.5: SDS-PAGE analysis of protein expression of ^{TmaPfu}DAH7PS from *E. coli* Chaperone 3 (DE3) cells after lysis in the debris (lane 1) and supernatant (lane 2).

proteins.[159] Therefore, the vector containing ^{TmaPfu}DAH7PS was transformed into *E. coli* BL21(DE3) pBB540/pBB542 (Chaperone 3) cells. These cells contain a number of chaperones, dnsK, dnaJ, grpE, clpB and groESL, which are thought to help improve protein solubility by allowing chaperone-assisted folding (Figure 6.4).

The co-expression of ^{TmaPfu}DAH7PS with the *E. coli* chaperonins resulted in soluble protein, as seen by SDS-PAGE analysis through protein of the appropriate molecular weight in the supernatant fraction. SDS-PAGE analysis of the supernatant and pellet fractions after cell lysis showed that approximately fifty percent of the total protein expressed was present in the soluble fraction (Figure 6.5), however this appeared to be a considerable improvement from expression in the *E. coli* BL21(DE3) cells (Figure 6.3). Therefore, purification of this enzyme continued. ^{TmaPfu}DAH7PS was purified using the same methods used for *Tma*DAH7PS, namely heat treatment followed by HIC and SEC (Chapter 2, Section 2.3).

6.4 Confirmation of Molecular Weight

The molecular mass of ^{TmaPfu}DAH7PS was 37068 Da as determined by electrospray ionisation mass spectrometry. This value is identical to the molecular weight calculated from the 333 amino acid ORF for ^{TmaPfu}DAH7PS.

6.5 Circular Dichroism Spectrophotometry

To assess protein folding of the chimeric protein, the secondary structure of ^{TmaPfu}DAH7PS was investigated using circular dichroism (CD) spectrophotometry. The spectrum recorded for ^{TmaPfu}DAH7PS (Figure 6.6) is similar to the closely-related KDO8PS enzymes, which have a similar fold,[48] and indicates a secondary structure typical of a $(\beta/\alpha)_8$ tertiary structure.[160] Furthermore, comparison of the CD spectra of ^{TmaPfu}DAH7PS with that of *Tma*DAH7PS suggests the chimeric protein is correctly folded.

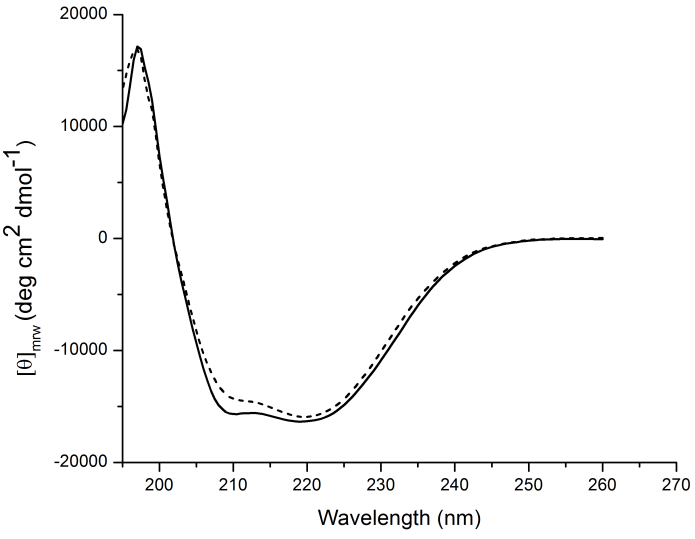


Figure 6.6: Comparison of the CD spectra of ^{TmaPfu}DAH7PS (solid line) and *Tma*DAH7PS (short dashes).

Table 6.1: Comparison of kinetic parameters for ^{TmaPfu}DAH7PS, *Tma*DAH7PS and *Pfu*DAH7PS enzymes.

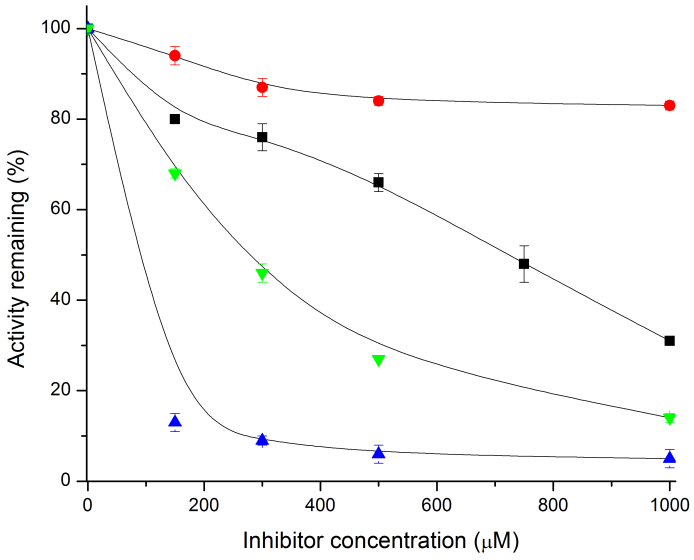
Enzyme	K_m^{PEP} (µM)	K_m^{E4P} (µM)	k_{cat} (s^{-1})	Temp. (°C)	Reference
^{TmaPfu}DAH7PS	46 ± 4	10 ± 1	2.41 ± 0.02	60	This study
*Pfu*DAH7PS	120 ± 20	28 ± 4	1.5 ± 0.1	60	Schofield[38]
*Tma*DAH7PS	4.85 ± 0.04	13 ± 1	11.7 ± 0.2	60	This study

6.6 Kinetic Parameters

^{TmaPfu}DAH7PS was found to be catalytically active, consistent with the observation showing the protein is correctly folded. The values of the kinetic parameters for freshly purified ^{TmaPfu}DAH7PS were determined in the same manner as for wild-type *Tma*DAH7PS by following the loss of PEP absorbance, spectrophotometrically, at 232 nm and 60 °C. The apparent K_m values for E4P and PEP were 10 ± 1 µM and 46 ± 4 µM, respectively, and k_{cat} of 2.41 ± 0.02 s^{-1}. These K_m values are similar to those for both *Pfu*DAH7PS and *Tma*-DAH7PS (Table 6.1). The k_{cat} is very close to that reported for *Pfu*DAH7PS, from which the catalytic barrel of ^{TmaPfu}DAH7PS is derived.[38] Furthermore, some variation of the k_{cat} value for *Pfu*DAH7PS has been previously noted (Schofield, unpublished data).

6.7 Feedback Inhibition

Two aromatic amino acids were tested as potential feedback inhibitors of ^{TmaPfu}DAH7PS, as described in Chapter 8, Section 8.6.4. Tyr and Phe at increasing concentrations were tested for their effect on enzyme activity. Tyr, and to a lesser extent Phe, were shown to have an inhibitory effect on the enzyme (Figure 6.7). The inhibition by Tyr is much smaller for ^{TmaPfu}DAH7PS than for wild-type *Tma*DAH7PS (Figure 6.7) and even at increased inhibitor concentrations (up to 1 mM), the enzyme may still not be maximally inhibited.



(a)

Figure 6.7: Effect of Tyr and Phe on the activity of ^{TmaPfu}DAH7PS in response to increasing concentrations of Tyr (■) and Phe (●) in comparison to the inhibition of *Tma*DAH7PS by Tyr (▲) and Phe (▼). Triplicate assays were performed and error bars represent the standard deviation.

6.8 Kinetic Parameters in the Presence of Tyr

To determine if the presence of Tyr has the same effect on catalysis as for *Tma*DAH7PS, kinetic parameters were determined for ^{TmaPfu}DAH7PS in the presence of Tyr (500 µM). The apparent K_m for PEP remained relatively unchanged, whereas the enzyme showed a reduced affinity for E4P (K_m^{PEP} of 42 ± 2 µM, K_m^{E4P} of 20 ± 2 µM) and slightly lower k_{cat} (1.6 ± 0.1 s^{-1}) when compared to the parameters determined for ^{TmaPfu}DAH7PS in the absence of Tyr (Table 6.1).

These results differ from the observed kinetic parameters for wild-type *Tma*DAH7PS determined in the presence of Tyr. In the presence of Tyr

(150 µM), *Tma*DAH7PS displayed a reduced affinity for both substrates, particularly for E4P, and a reduced k_{cat}. ^{TmaPfu}DAH7PS required a higher inhibitor concentration than wild-type *Tma*DAH7PS to affect catalysis and this reflects the heightened sensitivity of *Tma*DAH7PS towards inhibitors compared to ^{TmaPfu}DAH7PS.

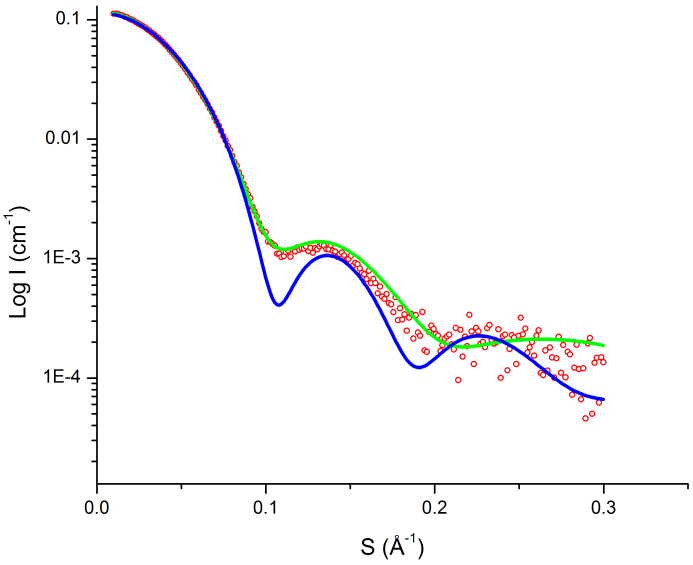
6.9 Small Angle X-ray Scattering

SAXS was employed to determine if the conformational change on binding of Tyr, seen in solution for *Tma*DAH7PS, also occurs for ^{TmaPfu}DAH7PS. Data were collected in the presence and absence of Tyr. In the absence of Tyr, the profile closely resembles the predicted profile for *Tma*DAH7PS (Figure 6.8a). This confirms that the protein adopts the same tetrameric structure and the same open conformation as seen for wild-type *Tma*DAH7PS. On treatment with 1 mM Tyr, the profile changed significantly, which is in line with the changes observed for wild-type *Tma*DAH7PS (Figure 6.8b).

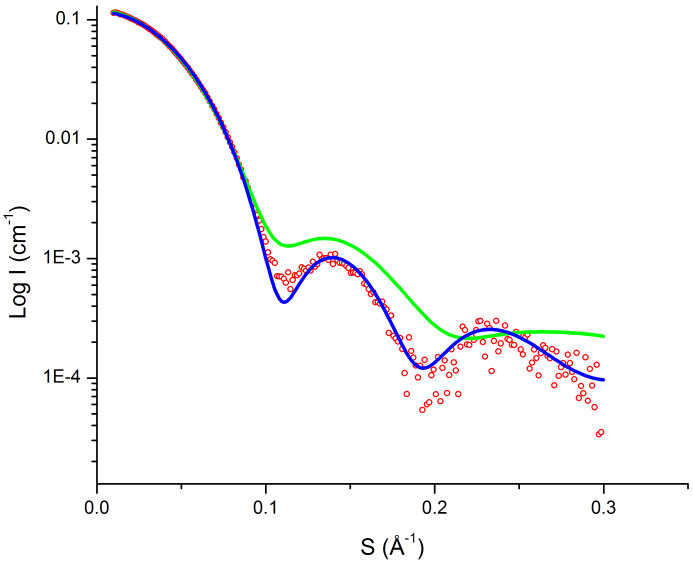
For the scattering profile of ^{TmaPfu}DAH7PS in the absence of Tyr, the radius of gyration (R_g), as determined by Guinier analysis, was 34.2 ± 0.2 Å, and as calculated using the indirect Fourier transform method from the pair distribution function ($P(r)$), 34.3 ± 0.1 Å (Figure 6.9a and 6.9c). The maximum dimension of the scattering particle (D_{max}) was approximately 100 Å, which is in close agreement with the Tyr-free open crystal structure of wild-type *Tma*DAH7PS (PDB code 1RZM) when analysed with CRYSOL (D_{max} of approximately 106 Å) with a χ^2_ν of 0.7.

In the presence of Tyr, the R_g determined from the Guinier analysis (Figure 6.9b) and $P(r)$ are similar (33.1 ± 0.1 Å and 32.8 ± 0.03 Å, respectively) and the D_{max} calculated from the $P(r)$ was approximately 90 Å (Figure 6.9d). Again, this is in close agreement with the CRYSOL analysis of the closed Tyr-bound *Tma*DAH7PS structure (D_{max} of approximately 95 Å) with a χ^2_ν of 1.4.

These data suggest that the engineered ^{TmaPfu}DAH7PS enzyme behaves



(a)



(b)

Figure 6.8: SAXS measurements of *TmaPfu*DAH7PS in the absence and presence of Tyr. Theoretical scattering coordinates were generated from crystallographic coordinates from both the open Tyr-free *Tma*DAH7PS, PDB code 1RZM (—) and Tyr-bound *Tma*DAH7PS structure PDB code 3PG9,(—) using CRYSOL. These theoretical plots have been compared to the experimentally determined scattering of *TmaPfu*DAH7PS (○) in the (a) absence and (b) presence of Tyr. Reduced chi-squared (χ^2_ν) for the fit to the calculated CRYSOL data for the open (PDB code 1RZM) and closed conformation (PDB code 3PG9) values are for (a) 0.70 and 5.12 and (b) 7.53 and 1.40 respectively.

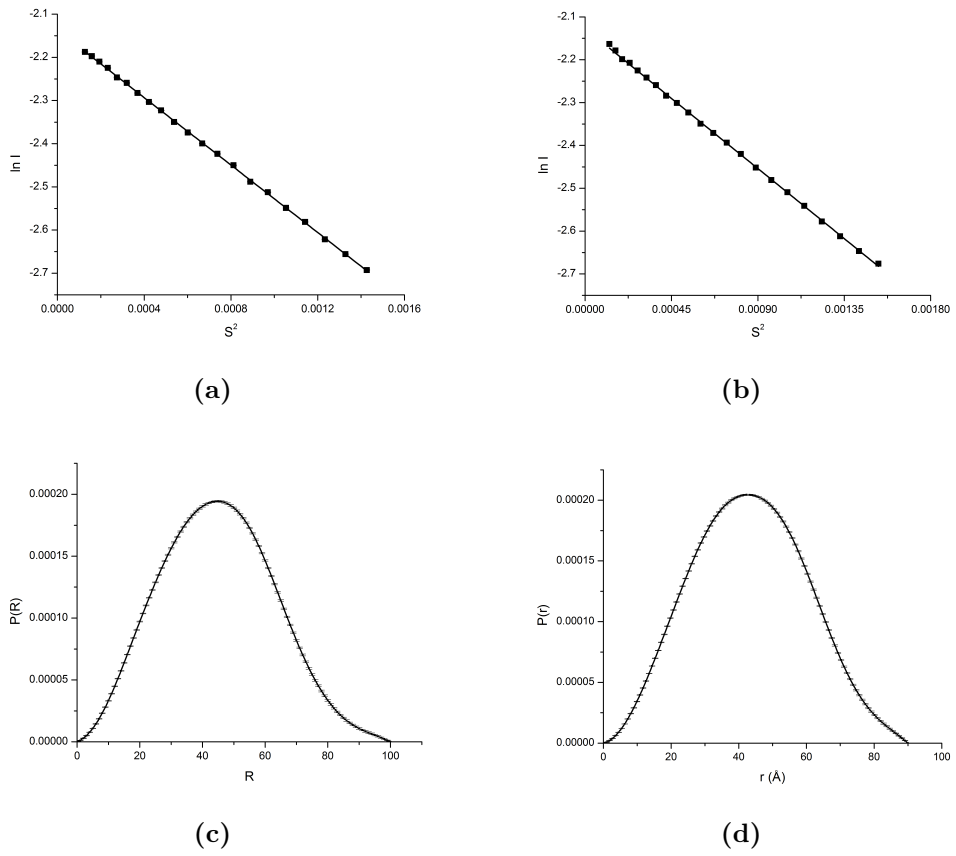
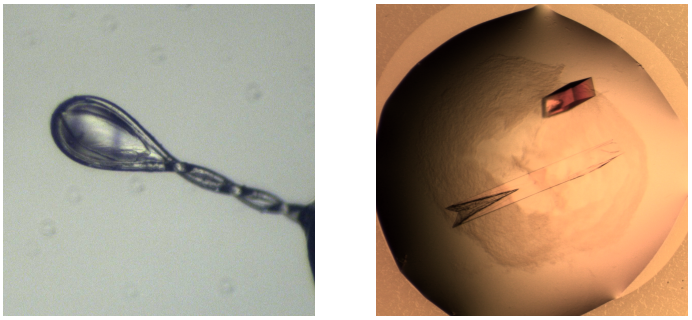


Figure 6.9: For ^{TmaPfu}DAH7PS data collected in the (a) absence and (b) presence of Tyr, Guinier plots were linear for $s \cdot R_g < 1.3$. $P(r)$ function calculated, using GNOM,[125] from ^{TmaPfu}DAH7PS SAXS data (c) without (d) with Tyr. Error bars indicate uncertainty in $P(r)$ propagated from the $I(s)$ vs s profile.

similarly to *Tma*DAH7PS in the presence of Tyr and adopts a more compact, globular structure. The SAXS and inhibition results for ^{TmaPfu}DAH7PS suggest that the fusion of the regulatory ACT domain from *Tma*DAH7PS to the catalytic barrel of *Pfu*DAH7PS transfers the property of allosteric control of catalytic function.



(a) **(b)**

Figure 6.10: The best diffracting crystals of ^{TmaPfu}DAH7PS from (a) C3 and (b) in-house crystal trials.

6.10 Crystallisation and Structure

6.10.1 Crystallisation Trials

Crystal trials were performed both in house and at the Collaborative Crystallisation Centre (C3) operated by CSIRO. Initial screens were set up at C3 using the sitting-drop vapour diffusion method at 281 K or 293 K, in the presence of Tyr.

The best diffracting crystal grown at C3 grew from a reservoir containing 0.2 M sodium chloride, 0.1 M Tris, pH 8 and 20 % (w/v) PEG 6000 (Figure 6.10a). In house crystal trials were performed using the sitting-drop vapour diffusion method and either JCSG+ or PACT Suite Crystal Screens, supplemented with Tyr and $MnSO_4$ and incubated at 293 K. The screen conditions that grew crystals were optimised by varying precipitant and salt concentration using the hanging-drop diffusion method. The best diffracting crystal resulting from in house trials (Figure 6.10b) grew from a reservoir containing 0.2 M CaCl, 0.1 M sodium acetate, pH 4.6, 300 µM $MnSO_4$, 0.02% (w/v) sodium azide and 1 mM Tyr.

6.10.2 Structure

Refinement of X-ray data collected from the crystal grown from the C3 screens (Table 6.2) indicates that ^{TmaPfu}DAH7PS adopts a closed conformation in the presence of Tyr (Figure 6.11). Electron density confirms the presence of Tyr bound in the same location as in wild-type *Tma*DAH7PS. As the resolution obtained for this structure is relatively low (3 Å), higher resolution data are being sought.

6.11 Discussion

The primary goal of protein engineering is to create a novel protein that possesses some improved or novel property. The construction of hybrid enzymes can be achieved by changing the amino acid sequence and can be approached using a variety of different techniques, including the modification of existing protein frameworks, random mutagenesis, or the less stochastic approach of combining known properties in existing proteins.

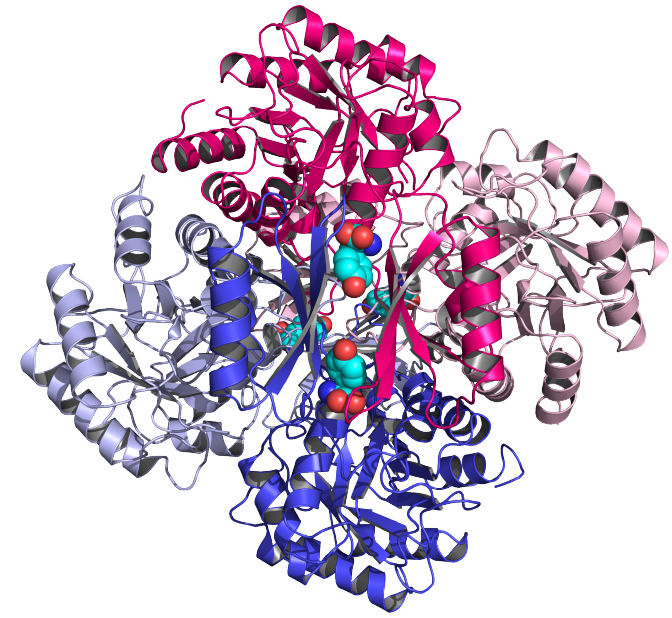
Point mutations and/or exchanges of secondary elements have been used to create chimeric or hybrid proteins with improved catalytic efficiency, stability, protein folding,[159] decreased susceptibility to proteases, exchanged or broadened substrate specificity,[161–167] regulated activity[168,169] or altered non-catalytic properties, such as thermostability.[170]

Although domain recruitment for the purpose of allosteric regulation is apparent in nature, and a number of bifunctional enzymes have been engineered by domain insertion,[171] the exchange of functional domains to engineer allosteric regulation has not been extensively explored.[152] A protein that couples two functions, such as an allosteric enzyme, or "molecular switch",[172] couples effector signals (input) to enzyme activity (output) and has potential for applications, such as biosensors,[171,173–176] novel biomaterials,[177] and modulators of gene transcription and cell signalling pathways.[178–180] Despite this vast potential, the myriad of design strategies and engineering

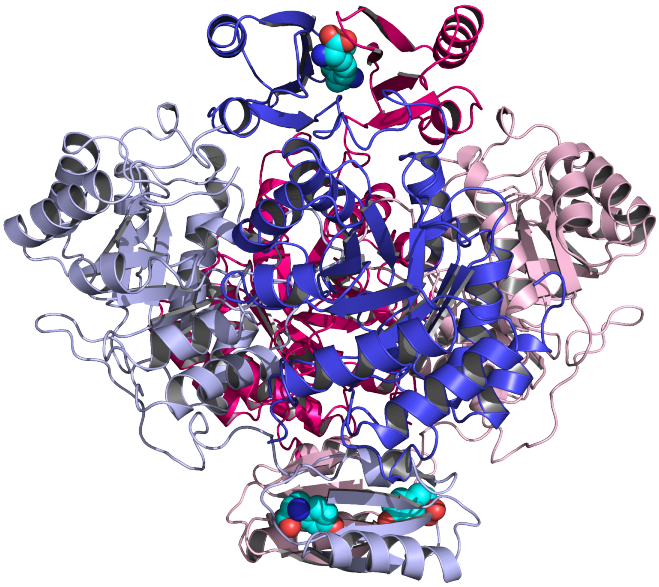
Table 6.2: Crystal parameters, data collection, and refinement statistics for ^{TmaPfu}DAH7PS in complex with Tyr

	^{TmaPfu}DAH7PS
Data collection	
Space group	$P2_12_12_1$
Cell dimensions	
a, b, c (Å)	76.9, 130.9, 138.1
α, β, γ (°)	90.0, 90.0, 90.0
Resolution range (Å)	37.94–3.00 (3.08–3.00)[a]
R_{merge}[b]	0.101 (0.581)
$I/\sigma(I)$	7.0 (1.3)
Completeness (%)	95.5 (97.6)
Redundancy	3.0 (3.0)
Refinement	
No. of reflections	80989
No. of unique reflections	27079
$R_{\text{work}}/R_{\text{free}}$[b]	0.195/0.276
No. of atoms	
Protein	9959
Ligand/ion	4
Water	23
B-factors (Å^2)	
Protein	61.57
Ligand/ion	–
RMSD	
Bond lengths (Å)	0.008
Bond angles (°)	1.228

[a]Values in parentheses are for the highest resolution shell.
[b]R_{merge}= $\sum h \sum i \, |Ii(h) - \langle I(h) \rangle| \, / \sum h \sum i Ii(h)$ and is provided by SCALA.
$r = \sum \| F_{obs}| - |F_{cal} \| \, / \sum |F_{obs}|$, where F_{obs} and F_{cal} are the observed and calculated structure-factor amplitudes, respectively for the working set (R_{work}) and test set (R_{free}) of reflections and are provided by REFMAC5.
Refinement parameters which have not yet been determined are denoted with dashes.



(a)



(b)

Figure 6.11: Preliminary structure of ^{TmaPfu}DAH7PS tetramer looking (a) down on and (b) side on to the tetramer with Tyr (shown as spheres with the carbon atoms coloured cyan) modelled into the ligand-binding site.

difficulties means there have been relatively few reported attempts to create allosteric enzymes.[179,181–185]

In the Iβ subfamily of DAH7PS, domain recruitment for the purpose of feedback regulation is clear from the inspection of sequence and structure. However, within this subfamily there are both regulated and unregulated forms of the enzyme. The creation of the chimeric protein ^{TmaPfu}DAH7PS, in which the regulatory ACT domain from one enzyme was fused onto an unregulated DAH7PS, resulted in a soluble, catalytically active and Tyr-inhibited enzyme. However, the catalytic activity of ^{TmaPfu}DAH7PS is not as sensitive to Tyr inhibition as wild-type *Tma*DAH7PS, the enzyme from which ^{TmaPfu}DAH7PS is modelled.

The ability of two ACT domains to interact appears to be the most important prerequisite for conferring allostery. It is easy to comprehend how this mechanism of regulation has evolved as it does not require any significant sequence or structural modifications to the core catalytic barrel. Additional to the contacts found between interacting regulatory domains, a number of contacts between the regulatory domain and catalytic barrel were highlighted on analysis of ligand-induced domain closure of *Tma*DAH7PS, specifically Arg277, Pro303, Glu304 and Asp309. Although these residues are not generally conserved within the barrel region of undecorated Iβ DAH7PS enzymes, all of these residues are conserved between *Pfu*DAH7PS and *Tma*-DAH7PS, and are therefore present in ^{TmaPfu}DAH7PS. Thus, it could be considered that transfer of allosteric regulation through fusion of a known regulatory element may have been assisted due to these critical contacts between the domains being present. The importance of these interactions to allostery is unknown. It appears that the ability and flexibility required for domain closure is a property of the domain itself but there are features of the catalytic barrel that could enhance this process.

The results obtained for ^{TmaPfu}DAH7PS support the hypothesis of domain recruitment for regulation and the associated implications of how the ACT domain confers allostery. Taken together with the results observed for ^{TmaPfu}DAH7PS, another study strengthens the potential for the transfer of

allosteric behaviour by domain recruitment. Random domain insertion of *E. coli* MBP into the penicillin-hydrolysing enzyme TEM-1 β-lactamase (BLA) was attempted in order to create an allosteric enzyme.[172] MBP was used to modulate the enzyme activity of BLA and this protein was chosen because of the large conformation change that occurs on ligand binding (a 35 ° closure and a −3.5 ° twist),[186] and because the conformational change observed for MBP has been exploited for the creation of molecular sensors.[175,176] Interestingly, one of the allosteric enzymes identified from the random library was essentially an end-to-end fusion of the two genes, MBP-BLA. This was an unexpected result as previous studies had not reported the switch behaviour of this construct.[182] However, an internal fusion produced a more ligand-sensitive enzyme and was deemed more successful.[172]

Unlike the MBP-BLA fusion protein in which only k_{cat} is affected by the presence of inhibitor (maltose),[172] ^{TmaPfu}DAH7PS behaves more like other allosteric enzymes where regulation is achieved through effects on K_{m} (binding) alone. SAXS and crystal data suggest the regulation mechanism of ^{TmaPfu}DAH7PS mimics that of *Tma*DAH7PS, where two ACT domains associate and bind Tyr between them, which physically gates substrate entry to the active site. Although a recurring motif, the ACT domain displays a multitude of variations in tertiary and quaternary architecture. Similarly, diversity of ligand-binding interactions is also noted in the small number of ligand-bound ACT domain structures solved. Despite these differences, it is obvious that the physical association between the ACT domains is a feature of their function.[130]

As more allosterically regulated enzymes are structurally characterised with ligands bound, other recurring domains similar to the ACT domain may be discovered, their mechanism for conferring allostery will be revealed and importance for regulation realised. Perhaps as more structure-function relationships are uncovered, other enzymes will reveal the importance of domain recruitment and large conformational changes associated with allosteric control.

Chapter 7

Summary of Thesis and Future Directions

This thesis describes functional and structural investigations of the type Iβ DAH7PS from *T. maritima*. The crystal structure of Tyr-bound *Tma*DAH7PS reported in this thesis is the first reported structure of an inhibitor-bound type Iβ DAH7PS enzyme.

The functional studies of *Tma*DAH7PS presented here expand upon initial examinations of this enzyme reported by others.[64] Kinetic parameters and inhibition data were broadly in line with results already published, confirming that *Tma*DAH7PS is a metal-dependent enzyme, which is strongly inhibited by Tyr and to a lesser extent by Phe.

7.1 The N-terminal βαββαβ Domain is Required for Allosteric Regulation

Truncation of the putative regulatory N-terminal domain, or ACT domain, of *Tma*DAH7PS resulted in a catalytically active enzyme. However, the truncated enzyme was insensitive to the inhibitors Tyr and Phe. Interestingly,

^{trunc}*Tma*DAH7PS was more catalytically active than wild-type *Tma*DAH7PS, suggesting the enzyme pays a catalytic penalty to accommodate the regulatory function. The crystal structure of ^{trunc}*Tma*DAH7PS showed the overall structure of the enzyme was not disturbed by removing the ACT domain. However, in contrast to the wild-type protein, which is tetrameric, this enzyme is dimeric in the crystal form. This change in quaternary structure was further investigated by SAXS where ^{trunc}*Tma*DAH7PS was found to be tetrameric, suggesting that the regulatory domain may stabilise quaternary structure. Further investigations into the quaternary structure of the enzyme and the importance of the regulatory domain for its stabilisation are required.

7.2 Mechanism of Allosteric Regulation Shown to Involve a Large Conformational Change

After establishing that the N-terminal domain of *Tma*DAH7PS is essential for allosteric regulation, the crystal structure of *Tma*DAH7PS in complex with Tyr was determined. Tyr was found to bind at the interface between two regulatory domains, formed from diagonally located monomers of the tetramer, revealing a major reorganisation of the regulatory domain with respect to the barrel relative to the unliganded enzyme. This significant conformational change observed in the crystal structure was corroborated by SAXS measurements collected in the presence and absence of Tyr. The closed conformation adopted by the protein on Tyr binding impedes substrate entry into the neighbouring barrel, revealing an unusual Tyr-mediated gating mechanism for regulation of this enzyme.

7.3 Hydrogen Bonding to Ser31 is Important for Inhibition by Tyr

Investigations into the important interactions between the inhibitor Tyr and enzyme showed that removal of the side chain of Ser31 substantially compromised inhibition by Tyr, and to a lesser extent by Phe. This finding suggests that the hydroxyl side chain of Ser31 is important for both the preferential inhibition by Tyr over Phe, and the inhibitory mechanism. Curiously, mutation of His29 (the hydrogen-bonding partner of Ser31) to Ala decreased sensitivity towards Tyr but increased inhibition by Phe. Exchange of the position of Ser31 and His29 showed that the relative positioning of these residues plays an important role in the inhibition of the enzyme. SAXS analysis showed that although the mutated enzymes were less sensitive to the inhibitory ligands, they still exhibited the conformational change observed for the wild-type enzyme in the presence of inhibitors. Residues that interact with the carboxylate or amino functionality of the bound-Tyr ligand or those that provide hydrogen-bonding contacts between two opposing ACT domains, namely Glu35, Arg36 and Arg46, could be targeted for mutation and investigated for their effect on Tyr binding.

7.4 Allosteric Regulation Can Be Transferred

Fusion of the ACT domain of *Tma*DAH7PS onto the undecorated, unregulated *Pfu*DAH7PS scaffold resulted in a catalytically active enzyme. Unsurprisingly, the enzyme had lower activity than its wild-type counterpart *Pfu*DAH7PS. However, the chimeric enzyme was inhibited by both Tyr and to a lesser extent, Phe. The conformational change observed by SAXS analysis for wild-type *Tma*DAH7PS was also observed for the fusion enzyme. This suggests that transfer of the property of allosteric regulation by fusion of an ACT domain is

indeed possible in line with the physical gating mechanism being principally a function of the regulatory domain. Although allostery was shown to be transferable, the inhibition was not as effective as for *Tma*DAH7PS, suggesting that features of the barrel-regulatory domain pairing require optimisation. This suggests that successful transfer is dependent on the presence of other factors that influence the transmission of the regulatory signal.

7.5 Future Experiments

Does Phe bind to *Tma*DAH7PS in the same site as Tyr?

Assuming Phe binds in the same position as Tyr, determination of the crystal structure of *Tma*DAH7PS in the presence of Phe may help explain why this inhibitor is not as effective as Tyr. It may also help explain the interactions that are important for Phe binding, and help elucidate what defines the specificity of the enzyme for Tyr over Phe.

What influence does the ACT domain have on quaternary structure?

Further investigations into the quaternary structure of *Tma*DAH7PS in solution are required, particularly to assess the different stabilities of the dimeric and tetrameric forms of the enzyme. Analytical ultracentrifugation (AUC) analysis of $^{\text{trunc}}$*Tma*DAH7PS may show whether a dimeric and/or tetrameric structure is adopted and whether quaternary structure is concentration-dependent. Futhermore, the tetramer interface of *Tma*DAH7PS could be disrupted by mutation to mimic the dimeric structure observed in the crystal form of $^{\text{trunc}}$*Tma*DAH7PS, and the mutated enzyme tested for sensitivity towards Tyr. This may uncover the interplay between quaternary structure and regulation.

Is the ACT domain transferable to other DAH7PS enzymes?

Fusion of the ACT domain onto another unregulated DAH7PS enzyme may show if allosteric regulation is generally a transferable property. Targets should include fusion to barrels where some of the interactions between the barrel and regulatory domain are unlikely to be formed. Although sequence analysis may assist in the identification of other unregulated DAH7PS enzymes, the lack of structural information and functional characterisation available somewhat limits the choice for a barrel scaffold to *Ape*DAH7PS. However, *Ape*DAH7PS shares every barrel contact to the ACT domain with *Tma*DAH7PS. Therefore, selecting a barrel template that shows variation in these residues may be more informative as to whether these contacts are essential for allosteric inhibition. In addition, mutations designed to disrupt the interactions between the ACT domain and barrel may also provide similar insights.

To assess the ability of the ACT domain to form a dimer without assistance from barrel residues, the ACT domain could be expressed and purified in the absence and presence of Tyr.

Concluding Remarks

Regulation is incredibly important for understanding enzymatic function and there are not many reported examples with all the information required to understand this phenomenon at a molecular level. *Tma*DAH7PS has illustrated a remarkable regulatory mechanism, and moreover, has been shown to have a mechanism that is readily transferable, allowing the evolution of complex regulatory systems to be unravelled.

Chapter 8

Experimental Procedures

8.1 General Methods

8.1.1 Amino Acid Sequence Alignments

Multiple amino acid sequence alignments were generated using ClustalW (http://www.ebi.ac.uk/clustalw/) or TCoffee (http://tcoffee.crg.cat/apps/tcoffee).

8.1.2 Protein Structure Figures

All protein structure figures were created using The PyMOL Molecular Graphics System (version 1.5, Schrödinger LLC).[187]

8.1.3 Purified Water

All water was purified by passage through a Millipore Milli-Q water system. This is referred to as Milli-Q water throughout this thesis. All Milli-Q water used for molecular biology was autoclaved prior to use.

8.1.4 pH Measurement

The pH of all solutions was measured at room temperature using a Denver UB-10 Ultrabasic pH meter with either a standard or micro-probe. The pH was adjusted using 1 M or 10 M HCl (or glacial acetic acid when required) or 1 M or 10 M NaOH.

8.1.5 Removal of Metal Ions from Solution

Metal ions were removed from solution by treatment with Chelex® 100 resin (Bio-Rad). The resin was added and left to stir overnight before the resin was removed by filtration (0.2 µm). The pH of all solutions was re-adjusted after Chelex® treatment.

8.1.6 Agarose Gel Electrophoresis

DNA fragments were separated on an agarose gel, prepared using a solution of 1 % (w/v) LE agarose (Seakem) in 50× Tris base-acetic acid-EDTA (TAE) buffer (2 M Tris-HCl, 1 M glacial acetic acid, 100 mM EDTA) heated in a microwave oven until dissolved and left to equilibrate at 60 °C before addition of SYBR® Safe DNA Gel Stain (Invitrogen). Samples were combined with 6× loading buffer (60 mM Tris-HCl, 60 mM EDTA, 0.02 % (w/v) orange G, 0.05 % (w/v) xylene cyanol FF, 60 % glycerol). Hand-cast gels were run in 1× TAE buffer using a Mini-sub cell GT (Bio-Rad) at 85 V until the dye front had migrated far enough to allow sufficient separation of the samples. Gels were visualised under UV (302 nm) and images were captured using a Molecular Imager® Gel Doc™ XR (Bio-Rad).

Alternatively, DNA samples were loaded on an E-Gel® 1.2 % precast agarose gel (Invitrogen) in a buffer-free system and run for 27 min. Separation could be monitored in real-time using the E-Gel® iBase™ Power System and E-Gel® Safe Imager™ Transilluminator at 480 nm.

For size comparison DNA fragments were run alongside a 1 kB Plus DNA ladder (Invitrogen).

8.1.7 LR Recombination Reaction

Sequenced plasmids were used to perform the Gateway® LR recombination reaction, a reaction that facilitates transfer of a Gateway® entry clone to a destination or expression vector. The LR Recombination Reaction consisted of approximately 100 ng of sequenced plasmid, approximately 150 ng of the appropriate destination vector and Tris-EDTA (TE) Buffer, pH 8 (10 mM Tris-HCl, pH 8.0, 1 mM EDTA) made up to a final volume of 8 µL. LR Clonase™ II enzyme mix (Invitrogen) was thawed on ice and gently mixed. 2 µL of the enzyme mix was then added, and the reaction mixture incubated at 25 °C for 1 h. The reaction was terminated by the addition of 1 µg of Proteinase K solution. The reaction mixture was then further incubated for 10 min at 37 °C.

8.1.8 Transformation of LR Recombination Reaction

The LR Recombination Reaction was used to transform competent *E. coli* One Shot® TOP10 Chemically Competent Cells (Invitrogen). A 50 µL aliquot of competent cells was defrosted and incubated on ice for 30 min with 1 µL of the reaction mixture. The cells were then heat-shocked for 30 s at 42 °C and immediately transferred to ice before 450 µL of room temperature super optimal broth (SOC) was added. The cells were incubated at 37 °C for 1 h before being spread on pre-warmed LB-Amp plates and left to grow overnight at 37 °C.

8.1.9 Transformation

Chemically competent cells (50 µL) were thawed on ice before the addition of 2 µL of either purified PCR product of extracted plasmid. The mixture

was left on ice for a further 30 min before heat-shocked at 42 ℃ for 30 s to 45 s. The cells were then replaced on ice for 2 min before the addition of 200 µL to 250 µL of SOC medium (2 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgSO_4$ and 20 mM glucose, filter sterilised) and outgrown at 37 ℃ for 1 h with shaking (200 rpm). An aliquot of transformed cells (75 µL to 200 µL) was spread directly onto a pre-warmed lysogeny broth (LB)-agar plate containing the appropriate antibiotics and left to grow overnight, inverted, at 37 ℃.

8.1.10 Chemically Competent Cells

All chemically competent cells used in this thesis were commercially prepared and supplied (Invitrogen or Novagen).

8.1.11 Plasmid Extraction and Isolation

Plasmids were extracted from 5 mL cultures, containing the appropriate antibiotic, and grown overnight at 37 ℃ using a High Pure Plasmid Isolation Kit (Roche). The concentration of purified plasmid was measured by absorption at 260 nm using a Nanodrop-1 000 spectrophotometer.

8.1.12 DNA Sequencing

DNA sequencing services were provided by Canterbury Sequencing and Genotyping based at the University of Canterbury. Sequencing was carried out using an Applied Biosystems 3130xl Genetic Analyzer and BigDye Terminator v3.1 (Applied Biosystems) used for the sequencing chemistry. Approximately 250 ng of purified double-stranded plasmid and 3.2 µM of primer were supplied for each sample.

8.1.13 Antibiotic Stocks

Stock solutions at 100 mg mL^{-1} of Amp and kanamycin (Kan), and 10 mg mL^{-1} of spectinomycin (Spec) were made up in Milli-Q water. 30 mg mL^{-1} chloramphenicol (Cam) was made up in analytical reagent (AR) ethanol. All solutions were filter sterilised (0.2 µm) into sterile micro-centrifuge tubes and stored at −80 ℃.

8.1.14 Glycerol Stocks

All strains of *E. coli* were grown overnight at 37 ℃, of which an aliquot was taken. Glycerol was then added to give a final concentration of 20 % (v/v), the culture was then thoroughly combined in a sterilised 1.7 mL micro-centrifuge tube, snap frozen in liquid nitrogen and stored at −80 ℃.

8.1.15 Media

All *E. coli* cultures were grown in LB (Lennox L) at 20 g L^{-1} (Invitrogen) made up with Milli-Q water and were sterilised by autoclaving at 121 ℃ at 15 psi for 20 min. Agar was added to the media used for agar plates (25 g L^{-1}) (Invitrogen) before sterilisation. Prior to use, the LB-agar solution was heated in a microwave oven until boiling, then left in a 60 ℃ oven for at least 1 h. The appropriate antibiotics were added immediately prior to pouring. SOC media was added to freshly transformed *E. coli* cells. Solutions consisted of 2 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$ and 20 mM glucose. All components were added to Milli-Q water except for 10 mM $MgCl_2$ and 20 mM glucose, which were filter sterilised and added immediately prior to use.

8.1.16 Centrifugation

All centrifugation was carried out using the following:

- Sorvall Evolution RC (Thermo Scientific)
- Heraeus Multifuge 1 S-R (Thermo Scientific)
- Heraeus Fresco 17 (Thermo Scientific)
- Minispin® Centrifuge (Eppendorf)

8.1.17 Growth of *E. coli* Cells

All 1 L cultures of cells were grown from a 50 mL overnight preculture of LB, inoculated from a glycerol stock scraping and containing the appropriate antibiotic. Cultures were grown in baffled flasks at 37 °C with shaking (180 rpm) until mid-logarithmic phase (OD_{600} of 0.4 AU to 0.8 AU).

8.1.18 Induction of Protein Expression

Expression of genes inserted in pT7-7 and pProEX-HT vectors are under control of a *lac* promoter and are thereby induced by lactose. IPTG is a synthetic analogue of lactose that triggers transcription of the lac operon, thereby protein expression, but is not metabolised by the cell as lactose is, which ensures a constant concentration of inductant is maintained. To induce protein expression, IPTG (Roche) was added to the cultures to a final concentration of 1 mM.

8.1.19 Harvesting of Cells

Cells were harvested approximately four hours after induction by centrifugation (14 000 *g*, 4 °C, 20 min or 12 000 *g*, 4 °C, 15 min). Cell pellets were stored at −80 °C until lysis.

8.1.20 Cell Lysis

Cell lysis was performed in one of two ways: Lysis by sonication was performed using an Omni-Ruptor 4000 Ultrasonic Homogenizer (Omni International) with either a titanium $\frac{3}{8}$ inch or $\frac{5}{32}$ inch probe. Typically, cell pellets were resuspended in 20 mL chilled lysis buffer, sonicated in a beaker surrounded with ice with 4–6 multiples of 5 min at 80 % and 30 % to 40 % pulsation.

The alternative method for lysis was the more gentle chemical detergent lysis using BugBuster® 10× Protein Extraction Reagent (Novagen). Cell pellets were resuspended in lysis buffer with a volume of BugBuster® equivalent to $1/20^{\text{th}}$ of the original bacterial culture volume pelleted and Nuclease (Novagen) (an engineered promiscuous endonuclease) was added per mL of BugBuster®. The cell suspension was shaken at 37 ℃ for at least 20 min.

Cell debris was removed by centrifugation (12 000 *g*, 4 ℃, 30 min or 24,000 *g*, 4 ℃, 30 min or 30 000 *g*, 4 ℃, 15 min).

The lysis buffer for all proteins consisted of 10 mM 1,3-(tris(hydroxymethyl)-methylamino)propane (BTP), pH 7.3, 200 mM KCl, 200 µM PEP and 1 mM EDTA.

8.1.21 Fast Protein Liquid Chromatography

Fast protein liquid chromatography (FPLC) was carried out using a Bio-Rad Biologic Protein Chromatography system or a ÄKTApurifier™ 10 (GE Healthcare) at either 4 ℃ or room temperature. All buffers and solvents were filtered prior to use using a 0.2 µm filter (Millipore). Protein samples were also filtered (2 µm) before being loaded onto a 50 mL Superloop™ (GE Healthcare). The eluant was fractionated in 2 mL fractions and collected in 96-well plates.

8.1.22 Hydrophobic-Interaction Chromatography

The supernatant fraction after heat treatment was pooled, ammonium sulfate added to a final concentration of 1 M, and loaded onto a SOURCE™ 15Phe® (GE Healthcare) column (8 mL), equilibrated in 10 mM BTP, pH 7.5, 1 mM EDTA and 1 M $(NH_4)_2SO_4$ (Buffer A) at room temperature. The unbound protein was removed with four column volumes of Buffer A and the bound protein was eluted with five column volumes of Buffer B (10 mM BTP, pH 7.5, 1 mM EDTA) at a flow rate of 2 mL min^{-1}. Fractions that eluted with Buffer B and those that showed UV absorbance at 280 nm were analysed by SDS-PAGE. Those fractions containing the protein of interest were pooled and concentrated using a 20 mL 10 kDa molecular weight cut-off (MWCO) device (Sartorius Stedim Biotech).

8.1.23 Size-Exclusion Chromatography

SEC using a Superdex S200 26/60 (GE Healthcare) or Superdex S200 HR 10/300 column (GE Healthcare) was used as a final purification step for all proteins discussed in this thesis. This step was used to separate proteins from any contaminating DNA and to exchange proteins into more suitable buffers for crystallisation trials.

Protein was injected onto the column under isocratic conditions with a flow of 0.2 mL min^{-1} to 2 mL min^{-1}. The SEC buffer used for all proteins consisted of BTP (10 mM, pH 7.5) supplemented with KCl (100 mM), PEP (200 µM) and EDTA (1 mM).

All buffers were made up using Milli-Q water. Fractions that gave an absorbance at 214 nm and/or 280 nm were analysed by SDS-PAGE.

8.1.24 Cleavage of N-terminal MBP and His_6-tag by rTEV

Fractions containing DAH7PS, as determined by SDS-PAGE were pooled, washed with loading buffer using a desalting column to remove any maltose and dithiothreitol (DTT) (1 mM) added. The protein was then incubated for up to 24 h at room temperature (approximately 22 ℃) with recombinant TEV (rTEV) protease (1 OD_{280} of TEV protease per 10 OD_{280} protein) (rTEV kindly donated by Dr Andrew Muscroft-Taylor). The completeness of cleavage was assessed by SDS-PAGE. The rTEV was successfully removed in the next step of purification by heat treatment or reapplication to HisTrap™ column (GE Healthcare).

8.1.25 Polyacrylamide Gel Electrophoresis

SDS-PAGE was performed using the method of Laemmli[188] with either a hand-cast gel made up of a 4 % stacking gel and 12 % (w/v) separating gel, run using a Mini Protean II cell (Bio-Rad) or a NuPAGE® Novex 10 % Bis Tris 1.0 mM precast protein gel (Invitrogen) run in a XCell *SureLock*® Mini-Cell (Invitrogen). Samples were not boiled before application to gel. Novex® Sharp Prestained Protein Standards (Invitrogen) or Low Range SDS-PAGE molecular weight standards (Bio-Rad) were used. A constant voltage of 200 V was applied across the electrodes until the dye front reached the bottom of the gel. After electrophoresis, gels were stained for protein using Coomassie Brilliant Blue R250 in 50 % (v/v) methanol, 10 % (v/v) acetic acid in water for approximately 30 min. To remove the excess dye, the gel was then detained for 30 min in a solution containing 40 % (v/v) methanol and 10 % (v/v) acetic acid.

8.1.26 Buffer Exchange and Concentration of Protein

Proteins were concentrated during and after purification using either a 20 mL or 70 mL 10 000 Da MWCO device (Sartorius Stedim Biotech). All filtration devices were washed before use with 2–3 equivalent volumes of Milli-Q water. Protein solutions were buffer-exchanged by dilution with the desired buffer and repeatedly concentrated.

8.1.27 Determination of Protein Concentration

Protein concentration was determined in one of two ways: Firstly, by the method of Bradford, 1976,[189] using bovine serum albumin (BSA) as a standard. Standards were made from a 1 mg mL^{-1} stock BSA solution. The assays were performed by thoroughly mixing 20 µL of protein or standard solution with 1 mL of QuickStart Bradford Dye Reagent (Bio-Rad) and the absorbance at 595 nm used to calculate concentration using Cary UV software.

Alternatively, the protein concentration was measured by UV absorbance at 280 nm using a Nanodrop ND-1 000 spectrophotometer (Thermo Scientific) and 2 µL sample of purified protein. The extinction coefficient value for each protein was determined from the amino acid sequence using ExPASy-ProtParam tool (http://web.expasy.org/protparam/). From this, absorbance was converted to concentration using the Beer-Lambert Law. Triplicate measurements were made and averaged.

8.1.28 Storage of Enzymes

Before storage, all purified enzymes were exchanged into buffers containing no EDTA before being split into aliquots no larger than 200 µL, snap frozen in liquid nitrogen and stored at −80 °C.

8.1.29 Measurement of Molecular Weight

The masses of purified proteins were measured by electrospray ionisation using either a Micromass LCT Classic (Waters) or Bruker maXis™ 3G (Bruker Daltonics). Proteins were diluted from stored concentrations to $1\,\mathrm{mg\,mL^{-1}}$ with Milli-Q water.

8.1.30 Standard Enzyme Assays and Kinetic Measurements

The assay system used for DAH7PS was based on the assay used by Schoner and Herrmann,[28] as modified by Schofield et al.[38] for use at 60 °C. The consumption of PEP was monitored at 232 nm (ϵ of $2.6 \times 10^3\,\mathrm{M^{-1}\,cm^{-1}}$ at 60 °C or ϵ of $2.8 \times 10^3\,\mathrm{M^{-1}\,cm^{-1}}$ at 25 °C) using a Varian Cary 100 UV Visible Spectrophotometer. Measurements were made using either 1 cm or 1 mm path length quartz cuvettes. Standard reaction mixtures contained PEP (approximately 150 µM), E4P (approximately 200 µM) and $MnSO_4$ (approximately 100 µM) in BTP buffer (50 mM, pH 7.3) and the enzyme to make a total volume of either 1 mL or 250 µL. The BTP buffer was treated with Chelex before use. PEP and E4P solutions were made up in the BTP buffer solution. $MnSO_4$ and amino acid solutions were made up in Milli-Q water that had been treated with Chelex. Assays were performed at either 25 °C or 60 °C.

Assays Performed at 60 °C

Reaction mixtures containing BTP buffer, $MnSO_4$ and PEP were incubated for seven minutes at 60 °C before the addition of enzyme. The reaction mixture was further incubated for two minutes before the reaction was initiated by addition of E4P. Duplicate or triplicate readings of all measurements were recorded and averaged.

Assays Performed at 25 °C

Reaction mixtures containing BTP buffer, $MnSO_4$ and PEP and E4P were incubated at 25 °C before the reaction was initiated by addition of enzyme. Duplicate or triplicate readings of all measurements were recorded and averaged. Initial rates of reaction were determined by a least squares fit of the initial rate data. One unit (1 U) of enzyme activity is defined as the loss of 1 µmol of PEP per minute at 60 °C or 25 °C. Specific activity is defined as the loss of 1 µmol of PEP per minute at 60 °C or 25 °C per mg of protein ($U\,mg^{-1}$). K_m and k_{cat} values were determined by fitting the data to the Michaelis-Menten equation using Grafit (Erithacus).

8.1.31 Determination of Substrate Concentration

The concentration of PEP and E4P was determined using the standard assay system. For example, for the concentration of E4P, a reaction mixture consisting of approximately 150 µM PEP, 100 µM $MnSO_4$, and enzyme is made up before being initiated with the addition of 5 µL of E4P stock solution (of unknown concentration) to give a final concentration of 1 mL. The reaction is allowed to go to completion and the difference in absorbance before initiation with E4P and the absorbance at completion of the reaction is measured. To convert absorbance into concentration the Beer-Lambert Law is applied ($A = \epsilon\ c\ l$, where $l = 1\,cm$ and ϵ of $2.8 \times 10^3\,M^{-1}\,cm^{-1}$ at 25 °C or ϵ of $2.6 \times 10^3\,M^{-1}\,cm^{-1}$ at 60 °C).

8.1.32 Crystallisation Trials

Screening for crystallisation conditions using sitting-drop vapour diffusion in 96-well plates (Hampton Research) at 293 K were performed by mixing 15 µL of protein solution with 15 µL of well solution. Initial screens were performed at the Collaborative Crystallisation Centre using the JCSG-plus and PACT-screens (Qiagen). The concentration for protein solutions used

for crystallisation trials were either 10 mg mL^{-1} (*Tma*DAH7PS) or 5 mg mL^{-1} ($^{\text{trunc}}$*Tma*DAH7PS).

Further manual screening was performed using both hanging-drop and sitting-drop vapour diffusion in 96- or 24-well plates (1 µL + 1 µL drops and 40 µL or 500 µL, respectively of appropriate well solution). Crystal trays were set up at room temperature and then transferred to 20 °C. Crystal trays and pipette tips were flushed with nitrogen, to remove any contaminating particles.

8.1.33 X-ray Data Collection

For X-ray data collection, crystals were transferred to reservoir solution containing 20 % (v/v) glycerol and flash-frozen in liquid nitrogen. Diffraction data were collected at the Australian Synchrotron using the MX2 beamline with a ADSC Quantum 315r detector (Area Detector Systems Corporation) at a temperature of 110 K. Data were collected by Dr Renwick Dobson.

8.1.34 Phasing and Model Refinement

Diffraction data were processed using MOSFLM[134] and SCALA.[135] Initial phase estimates were calculated by molecular replacement using PHASER.[190] The R_{free} set (5 %) was chosen from thin shells. Structural refinement was performed using REFMAC5[191] or PHENIX,[192] with iterative model building using COOT.[193] NCS was included in the initial rounds of refinement and waters were gradually added in the later rounds. TLS refinement was included in the later rounds of refinement, with groups chosen using the TLSMD server.[194] Structures were validated using the MolProbity server.[195]

8.1.35 Feedback Inhibition Studies

Solutions of either Phe (20 mM), Tyr (20 mM) or Trp (20 mM) made up in Chelex®-treated water were added to standard assay reaction mixtures to give a total reaction volume of either 250 µL or 1 mL. Assays were initiated by addition of E4P. All reactions were performed in duplicate or triplicate.

8.1.36 Small Angle X-ray Scattering Measurements

Measurements were performed at the Australian Synchrotron SAXS/WAXS beamline equipped with a Pilatus detector (1 M, 170 mm × 170 mm, effective pixel size, 172 × 172 µm). The wavelength of the X-rays was 1.0332 Å. The sample-detector distance was 1600 mm, which provided an s range of 0.0126–0.6015 $Å^{-1}$ (where s is the magnitude of the scattering vector, which is related to the scattering angle ($2\,\theta$) and the wavelength (λ) as follows: $s = (4\pi/\lambda)\sin\theta$). Scattering data were collected from enzymes (approximately 10 mg mL^{-1}) following elution from a size exclusion chromatography column (Superdex 200 5/150), pre-equilibrated (10 mM BTP pH 7.5, 100 mM KCl, 200 µM PEP) with or without amino acid (500 µM or 1 mM).

Data were collected from a 1.5 mm glass capillary at 27 °C at 2 s intervals. Scattering from glucose isomerase (10 mg mL^{-1}) were collected for calibration. Two-dimensional intensity plots from the peak of the size-exclusion chromatography run were radially averaged, normalised to sample transmission and background subtracted.

8.1.37 Small Angle X-ray Scattering Data Analysis

Scattered intensity (I) was plotted versus s. Extrapolation of the DAH7PS $I(s)$ profiles to zero angle ($I(0)$) and comparison with that of glucose isomerase standards indicated molecular mass. All samples were checked for an increase in intensity at low s (indicative of aggregation). Guinier plots were linear for $s \cdot R_g < 1.3$. 1D profiles were background subtracted and Guinier fits were made

using PRIMUS. Indirect Fourier transform was performed using GNOM[125] to yield the function $P(r)$, which gives both the relative probabilities of distances between scattering centers and the maximum dimension of the scattering particle ($D_{\max}$).

Theoretical scattering curves were generated from atomic coordinates and compared with experimental scattering curves using CRYSOL.[196] OLIGOMER,[197] which fits an experimental scattering curve from a multicomponent mixture of protein structures, was used to find the volume fractions of each component in the mixture. To assess and compare quality of the fits in CRYSOL or OLIGOMER, χ^2_ν or χ^2 values for the fits were compared, as outlined in Mills et al.[198]

8.2 Methods for Chapter Two

8.2.1 Amplification of the *Tma*DAH7PS Gene by PCR

The cloning of the gene for *Tma*DAH7PS into the pT7-7 expression vector was performed by Dr Mark Patchett prior to the start of this project. The DNA encoding the ORF of *Tma*DAH7PS was amplified from *T. maritima* MSB8 (DSM 3109) purified genomic DNA using primers designed specifically to introduce NdeI and BglII recognition sites. The resulting 1036 bp PCR product was ligated into pT7-7 previously digested with NdeI and BamHI, and transformed into *E. coli* BL21 Rosetta cells (Novagen).

8.2.2 Kinetic Parameters

The assays performed to determine the kinetic parameters for wild-type *Tma*-DAH7PS included 100 µM $MnSO_4$, 2 µL 1.75 mg mL^{-1} enzyme in 50 mM BTP pH 7.3 buffer. To determine K_m^{E4P}, the PEP concentration was fixed at 82 µM while the E4P concentration was varied (4 µM to 160 µM). For determination of K_m^{E4P}, E4P concentration was fixed at 161 µM while the PEP concentration was varied (3 µM to 82 µM).

To determine the kinetic parameters in the presence of Tyr, the same amount of enzyme was used, however 150 µM Tyr was added to each assay at the beginning of the incubation period. K_m^{E4P} in the presence of Tyr was determined by fixing the concentration of PEP at 125 µM while the E4P concentration was varied (31 µM to 310 µM). For determination of the K_m^{PEP}, E4P concentration was fixed at 160 µM while the PEP concentration was varied (11 µM to 165 µM).

8.2.3 Activity in Presence of Inhibitors

To determine the loss of activity in the presence of Tyr or Phe, reaction mixtures containing 145 µM PEP, 10 mM $MnSO_4$ in 50 mM BTP pH 7.3 buffer, 0 µM to 300 µM Tyr or Phe were incubated at 60 ℃ for seven minutes before the addition of 2 µL of 1 mg mL^{-1} enzyme. This mixture was incubated for a further two minutes before being initiated with 180 µM to 195 µM E4P to a final volume of 1 mL. Initial rates were compared to those of standard assay conditions to determine loss of activity. Triplicate assays were performed.

8.2.4 Metal Activation

The metal activation profile of wild-type *Tma*DAH7PS was determined using standard assays with 124 µM PEP and 161 µM E4P in 50 mM BTP containing 10 µM EDTA, pH 7.3 at 60 ℃ and 100 µM of $CoCl_2$, $CdCl_2$, $ZnCl_2$ or $MnSO_4$ made up in Chelex®-treated and filtered Milli-Q water. The enzyme (2 µL, 1.75 mg mL^{-1}) was incubated with the metal ion for two minutes prior to activation with E4P, to a final volume of 1 mL. Duplicate assays were performed.

8.2.5 Response to Allosteric Effectors

Solutions of either Phe (20 mM), Tyr (20 mM) or Trp (20 mM) made in water were added to standard assay reaction mixtures containing PEP (145.8 µM) and $MnSO_4$ (100 µM) in BTP buffer (50 mM, pH 7.3) to give a final aromatic amino acid concentration of 300 µM. The reaction mixture was incubated for seven minutes at 60 ℃ before the addition of enzyme (2 µL, 2.12 mg mL^{-1}), further incubation for 2 minutes and initiation with E4P (33 µM). Duplicate assays were performed.

8.2.6 Differential Scanning Calorimetry

DSC analysis was carried out on a Nano II DSC (TA Instruments). Protein solutions were dialysed against buffer containing 10 mM BTP, pH 7.5, 100 mM KCl, 200 µM PEP and 0.5 mM tris(2-carboxyethyl)phosphine (TCEP) and the concentration of protein adjusted to 1 mg mL^{-1}. The reference cell was filled with the final dialysis buffer. For sample analysis in the presence of Tyr, Tyr was added directly to both the enzyme and buffer to a final concentration of 1 mM. Samples were degassed before measurement and scanned from 20 °C to 130 °C at a rate of 1 °C min^{-1} under 3 atm of pressure.

8.3 Methods for Chapter Three

8.3.1 Amplification of $^{\text{trunc}}$*Tma*DAH7PS Open Reading Frame

The DNA corresponding to the ORF of $^{\text{trunc}}$*Tma*DAH7PS was amplified from wild-type plasmid using Phusion High Fidelity® DNA polymerase (New England Biolabs). Primers used for amplification were as follows:

$^{\text{trunc}}$*Tma*DAH7PS

For: 5' `CACCCTCGTTTCTCGTGAGTTCCATCC`

Rev: 5' `TCAATTCACCTTCACCCCCAGG`

The PCR mixture consisted of approximately 135 ng µL^{-1} of wild-type *Tma*DAH7PS plasmid DNA, 0.5 µM of both primers, 200 µM of each of the deoxyribonucleotide triphosphates (dNTPs), 1× Phusion® HF buffer and 0.01 U of Phusion® DNA polymerase in a total reaction volume of 50 µL. The thermo-cycling program consisted of an initial denaturation at 98 ℃ for 1 min and this was followed by 31 cycles of denaturation for 10 s, primer annealing at 63 ℃ for 30 s with extension at 72 ℃ for 30 s. The 807 base pair (bp) PCR product was subjected to Dpn1 treatment (0.5 µL per 25 µL PCR reaction for one hour at 37 ℃, purified by E-Gel® CloneWell™ agarose gel electrophoresis and further purified and concentrated using a DNA Clean and Concentrator Kit™ (Zymo Research Corporation).

8.3.2 Transformation into *E. coli* One Shot® TOP10 Chemically Competent Cells

The 807 bp PCR product encoding $^{\text{trunc}}$*Tma*DAH7PS was ligated into a pENTR™/TEV/dTOPO vector using Gateway® cloning technology. Fresh PCR product (approximately 150 ng) was combined with 1 µL of the salt

solution (1.2 M NaCl and 0.06 µM $MgCl_2$), and 1 µL of the TOPO® pENTR™/TEV/dTOPO vector (15 ng µL^{-1} to 20 ng µL^{-1} linearised plasmid DNA) in 50 % glycerol, 50 mM Tris-HCl, pH 7.4, 1 mM EDTA, 2 mM DTT, 0.1 % Triton X-100, 100 ng mL^{-1} bovine serum albumin (BSA), 30 µM bromophenol blue) to a final reaction volume of 6 µL, mixed gently and incubated for 5 min at room temperature.

The ligation mixture was used to transform competent *E. coli* One Shot® TOP10 Chemically Competent cells (Invitrogen). 50 µL of competent cells were thawed and incubated on ice for 15 min with 2 µL of ligation mixture. The cells were then heat shocked at 42 °C for 30 s and then returned to ice before 250 µL of room temperature SOC was added. The cells were incubated at 37 °C for 1 hr with shaking. The cells were plated onto prewarmed LB plates containing Kan and left to grow overnight at 37 °C.

Plasmid DNA was isolated from 5 mL cultures in LB containing Kan using a High Pure Plasmid Isolation Kit (Roche). Colony PCR was used to identify colonies containing inserts in the correct orientation and the correct number of bps. The PCR reaction consisted of 1× PCR buffer, 200 µM of each dNTP, 1.5 mM $MgCl_2$, 0.5 µM of both primers (one corresponding to the gene and the other to the plasmid), 1 U Taq DNA polymerase (Invitrogen) and approximately 50 ng of the plasmid extracted from each colony. Each PCR product that showed amplification of the approximate bp size of $^{\text{trunc}}$*Tma*-DAH7PS was DNA sequenced to identify a clone in which the insert PCR in the recombinant plasmid was identical to the expected ORF for $^{\text{trunc}}$*Tma*-DAH7PS.

8.3.3 LR Recombination Reaction and Vector Shuffle into pDEST566

A sequenced pENTR™/TEV/dTOPO $^{\text{trunc}}$*Tma*DAH7PS plasmid was used to perform the Gateway® LR recombination reaction and pDEST566, MBP-tagged, or pDEST15, His-tag destination vectors used.

8.3.4 Purification using MBP Affinity Chromatography

The supernatant collected after cell lysis and centrifugation was applied to a MBP-Trap (GE Healthcare) pre-equilibrated with Buffer A (10 mM BTP, pH 7.5, 1 mM EDTA, 100 mM KCl, 200 µM PEP) and eluted using Buffer C (Buffer A + 10 mM maltose). Fractions containing $^{\text{trunc}}$*Tma*DAH7PS were pooled and concentrated using a 20 mL 10 kDa MWCO device (Sartorius Stedim Biotech) and desalted. The MBP was cleaved from the $^{\text{trunc}}$*Tma*-DAH7PS using TEV protease (4 mg) and left stirring overnight at room temperature. The resulting supernatant was heat treated in covered Corex tubes at 60 °C for 40 min and left on the bench to cool to room temperature. Precipitated protein was removed by centrifugation at 24,000 g for 30 min at 4 °C. The supernatant was reapplied to MB-Trap (GE Healthcare) equilibrated with Buffer A, eluted in the flow through, concentrated using a 20 mL 10 kDa MWCO device (Vivascience), loaded to a 26/60 SEC column (GE Healthcare) and eluted using Buffer A at a flow rate of 2 mL min^{-1}. Fractions containing $^{\text{trunc}}$*Tma*DAH7PS were pooled, concentrated and reloaded to a pre-equilibrated (Buffer A) MBP-Trap and eluted in the flow through to remove any remaining cleaved MBP-tag. Again, fractions containing $^{\text{trunc}}$*Tma*-DAH7PS were pooled, concentrated using a 20 mL 10 kDa MWCO device (Vivascience).

8.3.5 Kinetic Parameters

The assays for determining the kinetic parameters for $^{\text{trunc}}$*Tma*DAH7PS included 100 µM $MnSO_4$, 2 µL 0.84 mg mL^{-1} enzyme in 50 mM BTP, pH 7.3 buffer. To determine the $K_{\text{m}}^{\text{E4P}}$, the PEP concentration was fixed at 80 µM while the E4P concentration was varied (20 µM to 310 µM). For the determination of $K_{\text{m}}^{\text{PEP}}$, E4P concentration was fixed at 310 µM while the PEP concentration was varied (10 µM to 165 µM).

8.3.6 Metal Activation

The metal activation profile of ^{trunc}*Tma*DAH7PS was determined using standard assays with 123 µM PEP and 309 µM E4P in 50 mM BTP pH 7.3, containing 20 µM EDTA at 60 °C and 100 µM of $CoCl_2$, $CdCl_2$, $ZnCl_2$ or $MnSO_4$ made up in Chelexed and filtered Milli-Q. The enzyme (2 µL 0.83 mg mL^{-1}) was incubated with the metal ion for 2 min prior to activation with E4P, to a final volume of 1 mL. Triplicate assays were performed.

8.3.7 Response to Allosteric Effectors

Reactions were carried out as per wild-type (refer to General Methods, Section 8.1.35), however 2 µL of 2.94 mg mL^{-1} purified enzyme was used.

8.3.8 Crystallisation Trials

Initial screens of ^{trunc}*Tma*DAH7PS were set up using the sitting-drop vapour diffusion method with droplets of protein solution (0.15 µL, 5 mg mL^{-1} in 10 mM BTP, 100 mM KCl, 200 µM PEP, pH 7.5) and reservoir solution (0.15 µL).

8.4 Methods for Chapter Four

8.4.1 Crystallisation Trials

For crystallisation trials of *Tma*DAH7PS with Tyr, initial screens were set up with droplets of 0.15 µL of protein solution (10 mg mL^{-1} in 10 mM BTP, pH 7.5, 100 mM KCl, 200 µM PEP and 1 mM Tyr) and 0.15 µL reservoir solution.

8.5 Methods for Chapter Five

8.5.1 Site-Directed Mutagenesis

The H29A, S31G and H29S/S31H mutations were performed using the QuikChange® Lightning Site-Directed Mutagenesis Kit (Stratagene), and pT7-7 *Tma*DAH7PS as the double stranded plasmid template. The single amino acid substitutions (either His-to-Ala, Ser-to-Gly or His-to-Ser and Ser-to-His) were introduced using synthetic oligonucleotide primers, each complementary to opposite strands of pT7-7 *Tma*DAH7PS, containing the desired mutation. For the primers shown below, codons representing the mutation are shown in lower-case letters.

*Tma*H29A

For: 5' GAGAGTTACAACTTGAAGTGTgccATTTCCAAAGGTCAGGAAAG

Rev: 5' CTTTCCTGACCTTTGGAAATggcACACTTCAAGTTGTAACTCTC

*Tma*S31G

For: 5' ACAACTTGAAGTGTCACATTggcAAAGGTCAGGAAAGAACGG

Rev: 5' CCGTTCTTTCCTGACCTTTgccAAAGGTCAGGAAAGAACGG

*Tma*H29S/S31H

For: 5' GGCCGAGAGTTACAACTTGAAGTGTagcATTcacAAAGGTCAGGAAAGAACGGTTATT

Rev: 5' AATAACCGTTCTTTCCTGACCTTTgtgAATgctACACTTCAAGTTGTAACTCTCGGCC

The procedure included three steps (as described in the QuikChange® manual): the primers with the desired mutation were extended during temperature cycling by a derivative of *Pfu*Ultra® high-fidelity (HF) DNA polymerase, generating a mutated plasmid containing staggered nicks. Following temperature cycling, the product was treated with DpnI to digest the parental DNA template. The nicked vector DNA containing the desired mutations was then transformed into One Shot® TOP10 competent cells (Invitrogen). Plasmid mini-prep DNA was then isolated from transformants using a High Pure Plasmid Isolation Kit (Roche) and the insert DNA sequenced to verify the

correct sequence before the plasmid-mutant construct was used to transform *E. coli* BL21(DE3) Star cells (Invitrogen).

8.5.2 Kinetic Parameters

Assays used to determine the kinetic parameters of all mutants of *Tma*-DAH7PS contained 100 µM $MnSO_4$, 2 µL of 0.52 mg mL^{-1} to 1.2 mg mL^{-1} enzyme in 50 mM BTP, pH 7.3 buffer. To determine the $K_{\mathrm{m}}^{\mathrm{E4P}}$, the PEP concentration was fixed at 80 µM (except for H29S/S31H for which PEP was fixed at 200 µM because of the higher PEP K_{m}), while the E4P concentration was varied (11 µM to 220 µM). For determination $K_{\mathrm{m}}^{\mathrm{PEP}}$, E4P concentration was fixed at 160 µM while the PEP concentration was varied (10 µM to 200 µM).

8.5.3 Response to Inhibitors

The assay mixtures to determine the effect of Tyr and Phe on the activity of the variants contained PEP (120 µM to 200 µM), E4P (200 µM to 220 µM and $MnSO_4$ (100 µM) in 50 mM BTP, pH 7.3. Reaction mixtures have either no amino acids or 0 mM to 1 mM Tyr or Phe. The reactions were performed as for standard assays and triplicate assays were performed for all reaction conditions.

8.6 Methods for Chapter Six

8.6.1 Ligation of ^{TmaPfu}DAH7PS into pDEST14

A synthetic gene encoding for the ORF of ^{TmaPfu}DAH7PS was ordered from GeneART® (Life Technologies). The synthetic gene had been optimised for expression in *E. coli*, inserted into the recombinational cloning entry vector pENTR221 and DNA sequenced before arrival. The ^{TmaPfu}DAH7PS gene was then ligated into the untagged Gateway expression vector pDEST14 using LR clonase II enzyme mix (Invitrogen) using Gateway Technology (as described in General Methods, Section 8.1.7). The reaction mixture was used to transform competent *E. coli* One Shot TOP10 and BL21(DE3) Star cells (Invitrogen).

8.6.2 Circular Dichroism Spectroscopy

CD spectroscopy experiments were performed using a JASCO J-815 Spectropolarimeter and spectra recorded from 260 nm to 190 nm using a 0.5 nm data pitch, 1 s response, 1 nm bandwidth at 25 °C. A 2 mL protein solution at a concentration of 0.03 mg mL^{-1} in a 3 mL quartz cuvette with a 1 cm path length was used. The buffer used was 10 mM phosphate, pH 7.0. A blank spectra of only buffer was recorded before the addition of protein.

The thermal denaturation was measured at 220 nm where a shift in the CD-spectrum indicative of the α-helix occurs. The CD was recorded every 2 °C while heating from 25 °C to 90 °C in a stopped cuvette, and the melt temperatures were calculated as the temperature of maximum inflection of the melting curve.

The Savitzky-Golay algorithm of the Lasco Spectra Manager™ (version 1.5) with a convolution width of 15 was used to smooth data.

8.6.3 Kinetic Parameters

To determine $K_{\mathrm{m}}^{\mathrm{E4P}}$, 100 µM $MnSO_4$, 240 µM PEP in 50 mM BTP, pH 7.3 was used with 2 µL of 2.03 mg mL^{-1} enzyme. E4P was varied between 28 µM and 440 µM.

For $K_{\mathrm{m}}^{\mathrm{PEP}}$, reaction mixtures contained 100 µM $MnSO_4$, 330 µM E4P, enzyme (2 µL, 2.03 mg mL^{-1}) in 50 mM BTP, pH 7.3, and between 10 µM and 400 µM PEP.

For the determination of $K_{\mathrm{m}}^{\mathrm{E4P}}$ in the presence of Tyr, 300 µM PEP, 100 µM $MnSO_4$ and 2 µL at 6.2 mg mL^{-1} in 50 mM BTP pH 7.3 and 500 µM Tyr were held constant while E4P concentration was varied between 11 µM and 220 µM.

66 µM E4P, 100 µM $MnSO_4$ and 2 µL at 6.2 mg mL^{-1} enzyme in 50 mM BTP pH 7.3 and 500 µM were used to determine $K_{\mathrm{m}}^{\mathrm{PEP}}$ in the presence of Tyr, whereas the concentration of PEP was varied between 40 µM and 400 µM.

8.6.4 Response to Allosteric Effectors

The response of the catalytic activity of ^{TmaPfu}DAH7PS was determined in the presence of 300 µM PEP, 100 µM $MnSO_4$, 110 µM E4P and 2 µL 6.2 mg mL^{-1} enzyme in 50 mM BTP, pH 7.3 and 0 mM to 1 mM of either Tyr or Phe. Triplicate assays were performed and averaged.

Appendices

Appendix A

Chapter Two Michaelis-Menten Plots For Determination of K_{m} Values Calculated for *Tma*DAH7PS

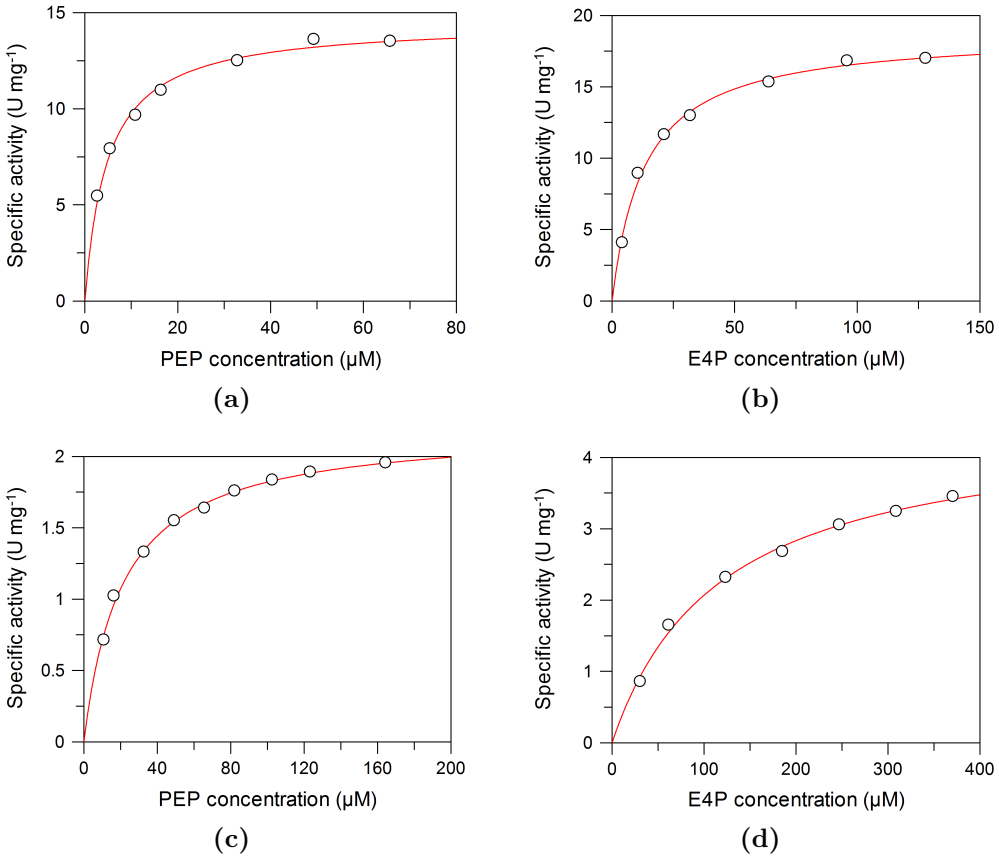


Figure A.1: Michaelis-Menten plots used for determination of the K_{m} values of both (a) PEP and (b) E4P for *Tma*DAH7PS, and (c) and (d), Michaelis-Menten plots used to ascertain the values for PEP and E4P in the presence of Tyr, respectively. K_{m} and k_{cat} values were determined by fitting the data to the Michaelis-Menten equation using Grafit (Erithacus Software).

Appendix B

Chapter Three Michaelis-Menten Plots For Determination of K_{m} Values Calculated for $^{\mathrm{trunc}}Tma$DAH7PS

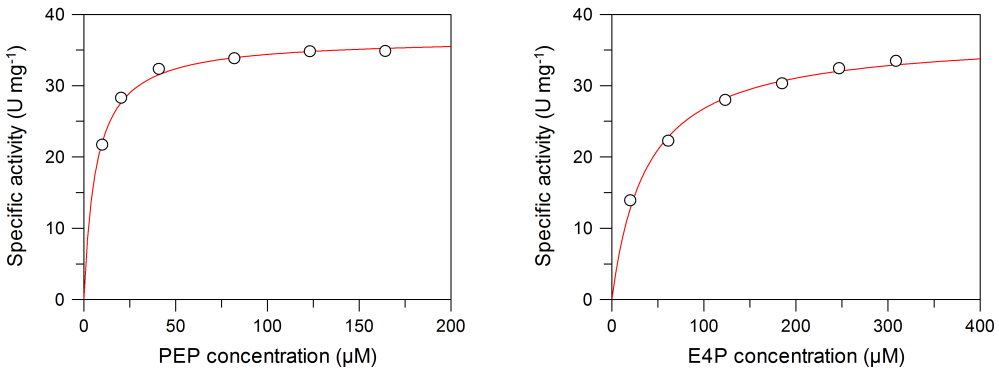


Figure B.1: Michaelis-Menten plots used for determination of the K_{m} values of both (a) PEP and (b) E4P for $^{\mathrm{trunc}}$*Tma*DAH7PS. K_{m} and k_{cat} values were determined by fitting the data to the Michaelis-Menten equation using Grafit (Erithacus Software).

Appendix C

Chapter Five Guinier Plots and $P(r)$ Functions Calculated From *Tma*DAH7PS Mutant SAXS Data

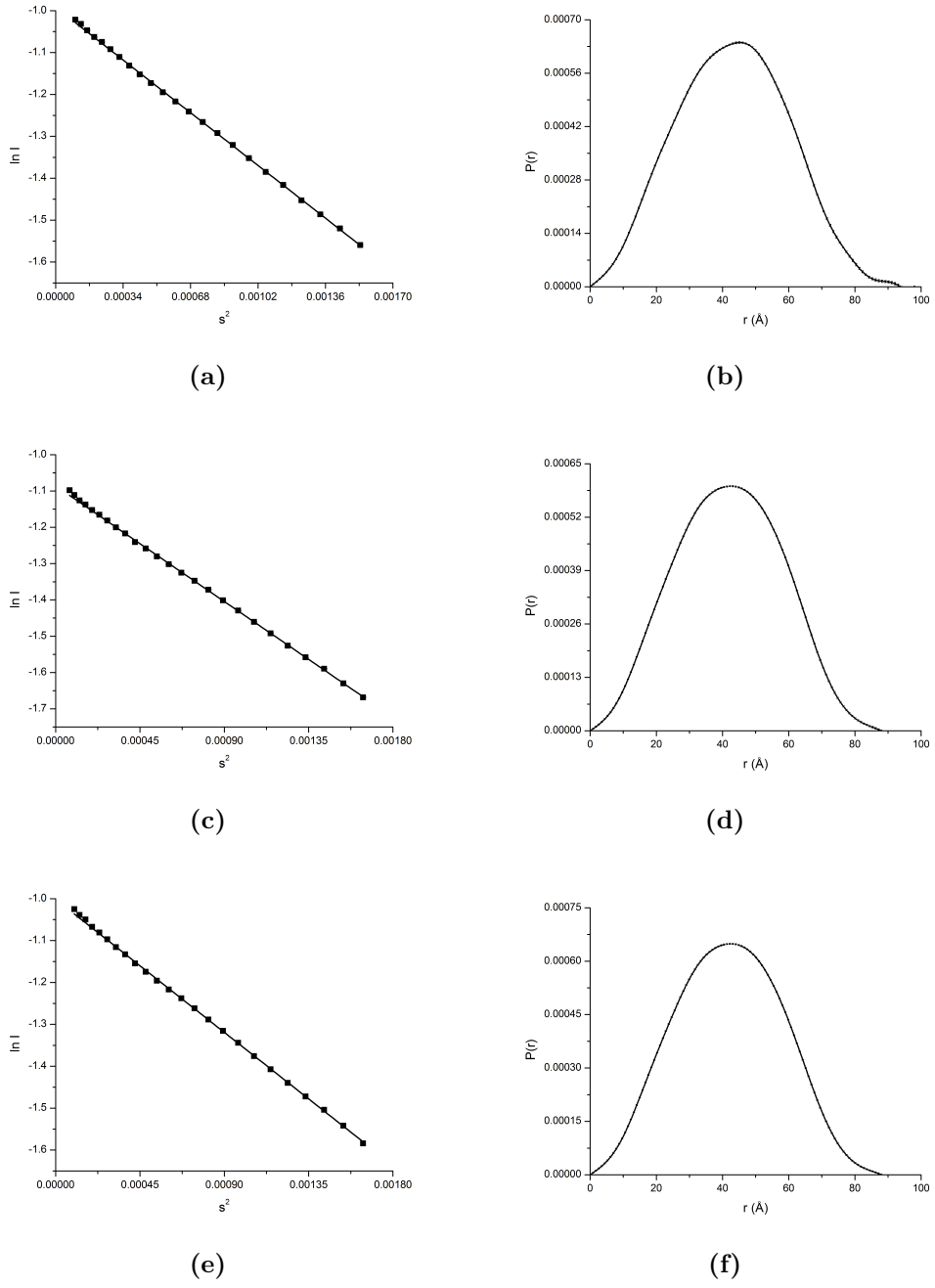


Figure C.1: Guinier plots and $P(r)$ functions calculated using GNOM from *Tma*DAH7PS SAXS data in the (a and b) absence of inhibitor, presence of (c and d) 1 mM Tyr and (e and f) 1 mM Phe. Uncertainty in $P(r)$ propagated from the $I(s)$ vs s profile is indicated by error bars.

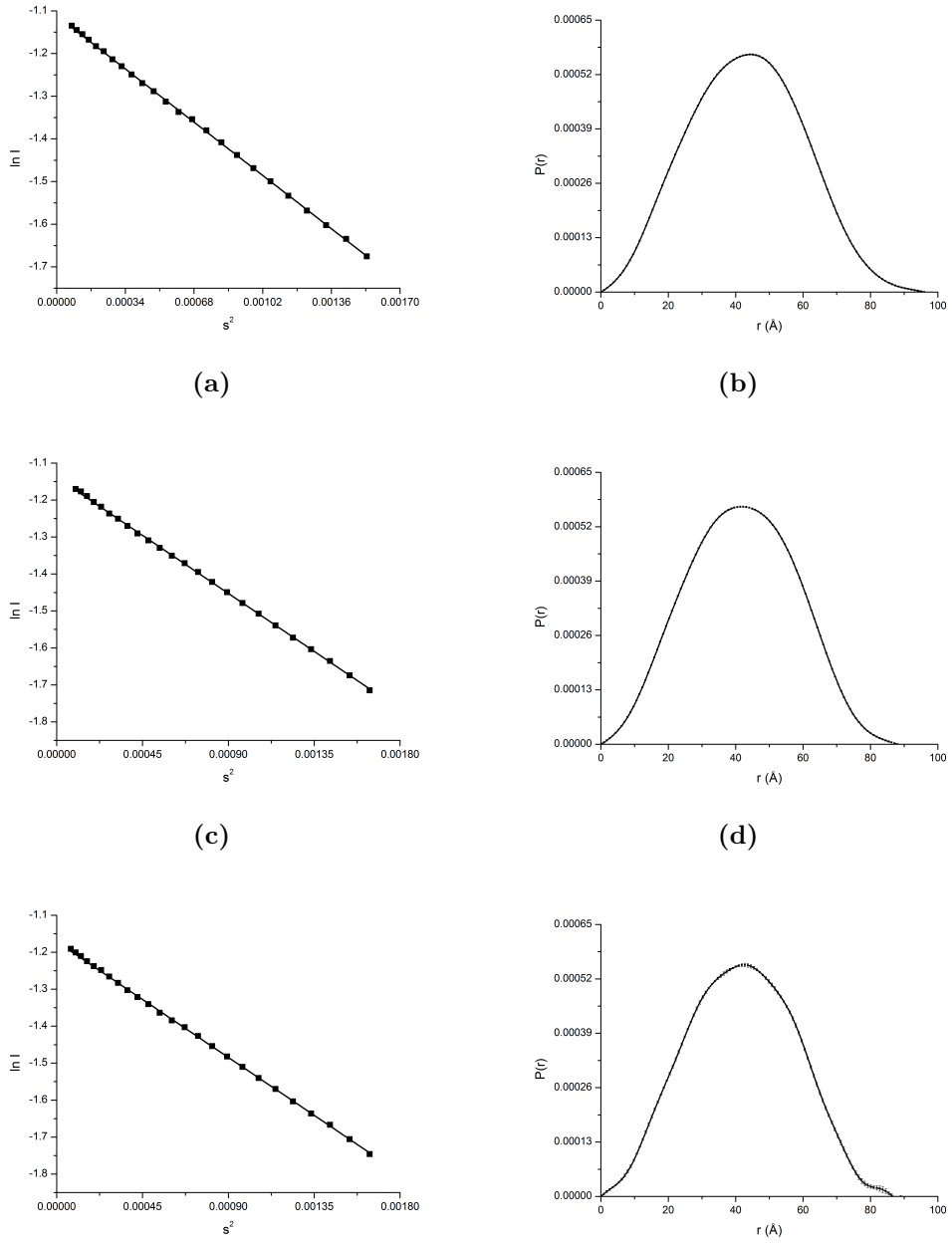


Figure C.2: Guinier plots and $P(r)$ functions calculated using GNOM from *Tma*H29A SAXS data in the (a and b) absence of inhibitor, presence of (c and d) 1 mM Tyr and (e and f) 1 mM Phe. Uncertainty in $P(r)$ propagated from the I(s) vs s profile is indicated by error bars.

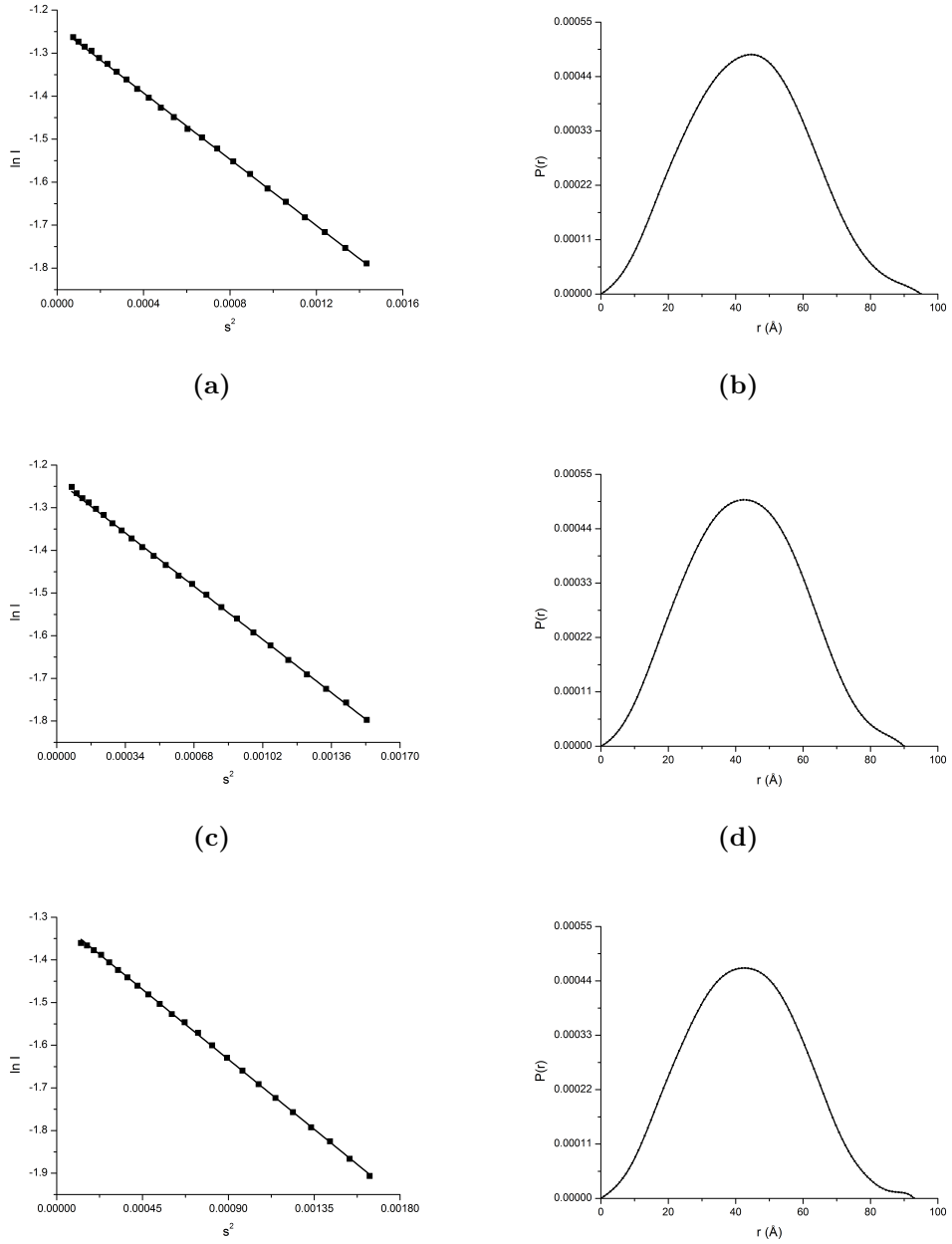


Figure C.3: Guinier plots and $P(r)$ functions calculated using GNOM from *Tma*S31G SAXS data in the (a and b) absence of inhibitor, presence of (c and d) 1 mM Tyr and (e and f) 1 mM Phe. Uncertainty in $P(r)$ propagated from the I(s) vs s profile is indicated by error bars.

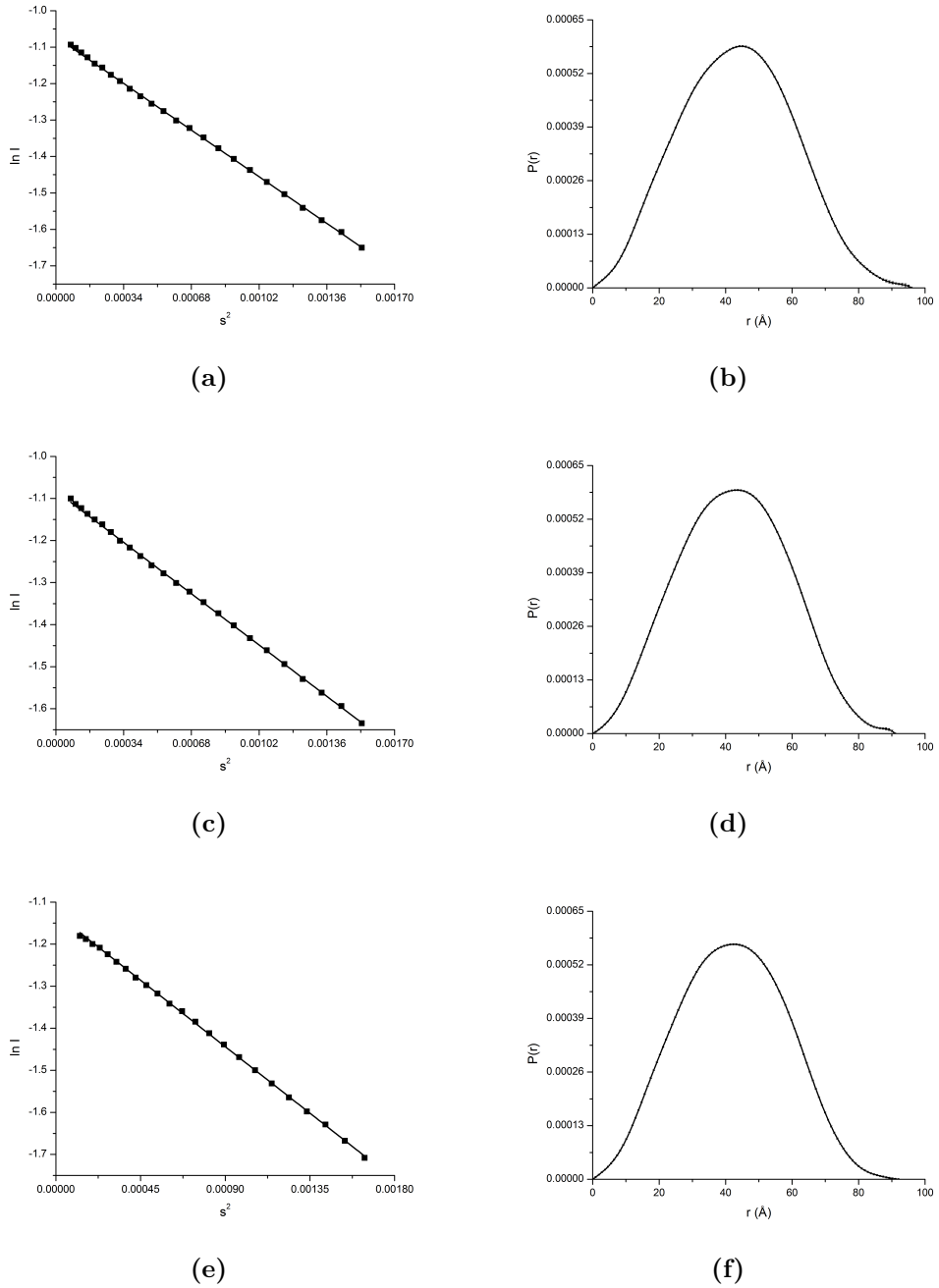


Figure C.4: Guinier plots and $P(r)$ functions calculated using GNOM from *Tma*H29SS31H SAXS data in the (a and b) absence of inhibitor, presence of (c and d) 1 mM Tyr and (e and f) 1 mM Phe. Uncertainty in $P(r)$ propogated from the I(s) vs s profile is indicated by error bars.

Appendix D

Chapter Six Codon Optimised Nucleotide Sequence of *TmaPfu*DAH7PS

```
ATG ATT GTT GTT CTG AAA CCG GGT AGC ACC GAA GAA GAT ATT CGT
AAA GTT GTT AAA CTG GCC GAG AGC TAT AAT CTG AAA TGC CAT ATT
TCA AAA GGC CAA GAA CGT ACC GTG ATT GGC ATT ATT GGT GAT GAT
CGT TAT GTT GTG GCC GAT AAA TTT GAA AGC CTG GAT TGT GTT GAA
AGC GTT GTT CGT GTG CTG AAA CCG TAT AAA CTG GTT AGC CGT GAA
TTT CAT CCG GAA GAT ACC GTT ATT GAT CTG GGT GAT GTG AAA ATT
GGC AAC GGC TAT TTT ACC ATT ATT GCA GGT CCG TGT AGC ATT GAA
AGC CGT GAT CAG ATT ATG AAA GTT GCA GAA TTT CTG GCC GAA GTG
GGT ATT AAA GTT CTG CGT GGT GGT GCA TTT AAA CCG CGT ACC AGC
CCG TAT AGC TTT CAG GGT TAT GGT GAA AAA GCA CTG CGT TGG ATG
CGT GAA GCA GCA GAT GAA TAT GGT CTG GTT ACC GTT ACC GAA GTT
ATG GAT ACC CGT CAT GTT GAA CTG GTT GCA AAA TAT AGC GAT ATT
CTG CAG ATT GGT GCA CGC AAT AGC CAG AAT TTT GAA CTG CTG AAA
GAA GTG GGC AAA GTT GAA AAT CCG GTT CTG CTG AAA CGT GGT ATG
GGT AAT ACC ATT CAA GAA CTG CTG TAT AGC GCA GAA TAT ATC ATG
GCA CAG GGT AAC GAA AAT GTG ATT CTG TGT GAA CGT GGT ATT CGT
ACC TTT GAA ACC GCA ACC CGT TTT ACC CTG GAT ATT TCT GCA GTT
CCG GTT GTT AAA GAA CTG AGC CAT CTG CCG ATT ATT GTT GAT CCG
AGC CAT CCG GCA GGT CGT CGT AGC CTG GTT ATT CCG CTG GCA AAA
GCA GCA TAT GCA ATT GGT GCA GAT GGT ATT ATG GTT GAA GTG CAT
CCG GAA CCG GAA AAA GCC CTG AGC GAT AGC CAG CAG CAG CTG ACC
TTT GAT GAT TTT CTG CAG CTG CTG AAA GAG CTG GAA GCA CTG GGT
TGG AAA GGT TAA ACC CAG CTT TCT TGT ACA AAG TGG TGG TAC CCT
GGA GCA CAA GAC TGG CCT CAT GGG CCT TCC GCT CAC TGC CCG
```

Figure D.1: Nucleotide sequence of the codon optimised ^{TmaPfu}DAH7PS.

Appendix E

Publication Resulting From Thesis Research

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 286, NO. 12, pp. 10216–10224, March 25, 2011

Tyrosine Latching of a Regulatory Gate Affords Allosteric Control of Aromatic Amino Acid Biosynthesis*□s

Received for publication, December 6, 2010, and in revised form, January 28, 2011 Published, JBC Papers in Press, January 30, 2011, DOI 10.1074/jbc.M110.209924

Penelope J. Cross‡§, **Renwick C. J. Dobson**‡¶_, **Mark L. Patchett****, **and Emily J. Parker**‡§1
*From the ‡Biomolecular Interaction Centre, University of Canterbury, Christchurch 8040, New Zealand, the §Department of Chemistry, University of Canterbury, Christchurch 8140, New Zealand, the ¶School of Biological Sciences, University of Canterbury, Christchurch 8140, New Zealand, the _Bio21 Molecular Science and Biotechnology Institute, Department of Biochemistry and Molecular Biology, University of Melbourne, Parkville, Victoria 3010, Australia, and the **Institute of Molecular Biosciences, Massey University, Palmerston North 4442, New Zealand*

The first step of the shikimate pathway for aromatic amino acid biosynthesis is catalyzed by 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (DAH7PS). *Thermotoga maritima* DAH7PS (*Tma*DAH7PS) is tetrameric, with monomer units comprised of a core catalytic (_/_)8 barrel and an N-terminal domain. This enzyme is inhibited strongly by tyrosine and to a lesser extent by the presence of phenylalanine. A truncated mutant of *Tma*DAH7PS lacking the N-terminal domain was catalytically more active and completely insensitive to tyrosine and phenylalanine, consistent with a role for this domain in allosteric inhibition. The structure of this protein was determined to 2.0 A˚ . In contrast to the wild-type enzyme, this enzyme is dimeric. Wild-type *Tma*DAH7PS was co-crystallized with tyrosine, and the structure of this complex was determined to a resolution of 2.35 A˚ . Tyrosine was found to bind at the interface between two regulatory N-terminal domains, formed from diagonally located monomers of the tetramer, revealing a major reorganization of the regulatory domain with respect to the barrel relative to unliganded enzyme. This significant conformational rearrangement observed in the crystal structures was also clearly evident from small angle X-ray scattering measurements recorded in the presence and absence of tyrosine. The closed conformation adopted by the protein on tyrosine binding impedes substrate entry into the neighboring barrel, revealing an unusual tyrosine-controlled gating mechanism for allosteric control of this enzyme.

References

[1] Davis, B. (1955) Intermediates in amino acid biosynthesis. *Adv. Enzymol. Relat. Subj. Biochem. 16*, 247–312.

[2] Sprinson, D. (1960) The biosynthesis of aromatic compounds from D-glucose. *Adv. Carbohydr. Chem. 15*, 235–270.

[3] Bentley, R., and Haslam, E. (1990) The shikimate pathway - a metabolic tree with many branches. *Crit. Rev. Biochem. Mol. Biol. 25*, 307–384.

[4] Roberts, F., Roberts, C., Johnson, J., Kyle, D., Krell, T., Coggins, J., Coombs, G., Milhous, W., Tzipori, S., Ferguson, D., Chakrabarti, D., and McLeod, R. (1998) Evidence for the shikimate pathway in apicomplexan parasites. *Nature 393*, 801–805.

[5] Herrmann, K., and Weaver, L. (1999) The shikimate pathway. *Annu. Rev. Plant Phys. 50*, 473–503.

[6] Steinrucken, H., and Amrhein, N. (1980) The herbicide glyphosate is a potent inhibitor of 5-enolpyruvyl-shikimic-acid 3-phosphate synthase. *Biochem. Biophys. Res. Commun. 94*, 1207–1212.

[7] Schönbrunn, E., Eschenburg, S., Shuttleworth, W., Schloss, J., Amrhein, N., Evans, J., and Kabsch, W. (2001) Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate 3-phosphate synthase in atomic detail. *Proc. Natl. Acad. Sci. U. S. A. 98*, 1376–1380.

[8] Eykmann, J. (1885) *Recl. Trav. Chim. Pay. B. 4*, 32.

[9] Haslam, E. *The shikimate pathway*; Butterworth & Co (Publishers) Ltd, London, 1974.

[10] Davis, B. D. (1951) Aromatic biosynthesis. *J. Biol. Chem. 191*, 315–326.

[11] Davis, B. D. (1952) Aromatic biosynthesis. V. Antagonism between shikimic acid and its precursor, 5-dehydroshikimic acid. *J. Bacteriol. 64*, 749–763.

[12] Weiss, U., Davis, B., and Mingioli, E. (1953) Aromatic biosynthesis. X. Identification of an early precursor as 5-dehydroquinic acid. *J. Am. Chem. Soc. 75*, 5572–5576.

[13] Salamon, I., and Davis, B. (1953) Aromatic biosynthesis. IX. The isolation of a precursor of shikimic acid. *J. Am. Chem. Soc. 75*, 5567–5571.

[14] Gibson, F., and Jackman, L. (1963) Structure of chorismic acid, a new intermediate in aromatic biosynthesis. *Nature 198*, 388–389.

[15] Davis, B., and Mingiolo, E. (1953) Aromatic biosynthesis. XII. Accumulation of two derivatives of shikimic acid by bacterial mutants. *J. Bacteriol. 66*, 129–136.

[16] Coggins, J., Boocock, M., Chaudhuri, S., Lambert, J., Lumsden, J., Nimmo, G., and Smith, D. (1987) The arom multifunctional enzyme from *Neurospora crassa*. *Method. Enzymol. 142*, 325–342.

[17] Srinivasan, P., Shigeura, H., Sprecher, M., Sprinson, D., and Davis, B. (1956) Biosynthesis of shikimic acid from D-glucose. *J. Biol. Chem. 220*, 477–497.

[18] Srinivasan, P., Sprinson, D., Kalan, E., and Davis, B. (1956) Enzymatic conversion of sedoheptulose-1,7-diphosphate to shikimic acid. *J. Biol. Chem. 223*, 913–920.

[19] Herrmann, K., Zhao, J., Pinto, a. W. L., J.E.B.P., and Henstrand, J. (1991) Regulation of carbon flow into the shikimate pathway. *Curr. Top. Plant Physiol. 7*, 12–18.

[20] Ogino, T., Garner, C., Markley, J., and Herrmann, K. (1982) Biosynthesis of aromatic compounds: ^{13}C NMR spectroscopy of whole *Escherichia coli* cells. *Proc. Natl. Acad. Sci. U. S. A. 79*, 5828–5832.

[21] Jensen, R., and Nester, E. (1965) Regulatory significance of intermediary metabolites – Control of aromatic biosynthesis by feedback inhibition in *Bacillus subtilis*. *J. Mol. Biol. 12*, 468–481.

[22] Nester, E., and Jensen, R. (1966) Control of aromatic acid biosynthesis in *Bacillus subtilis* – Sequential feedback inhibition. *J. Bacteriol. 91*, 1594–1598.

[23] Wu, J., Sheflyan, G., and Woodard, R. (2005) *Bacillus subtilis* 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase revisited: resolution of two long-standing enigmas. *Biochem. J. 390*, 583–590.

[24] Jensen, R., Xie, G., Calhoun, D., and Bonner, C. (2002) The correct phylogenetic relationship of KdsA (3-deoxy-D-*manno*-octulosonate 8-phosphate synthase) with one of two independently evolved classes of AroA (3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase). *J. Mol. Evol. 54*, 416–423.

[25] Shumilin, I., Kretsinger, R., and Bauerle, R. (1999) Crystal structure of phenylalanine-regulated 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*. *Struct. Fold. Des. 7*, 865–875.

[26] Wagner, T., Shumilin, I., Bauerle, R., and Kretsinger, R. (2000) Structure of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*: Comparison of the Mn(2+)*2-phosphoglycolate and the Pb(2+)*2-phosphoenolpyruvate complexes and implications for catalysis. *J. Mol. Biol. 301*, 389–399.

[27] Shumilin, I., Bauerle, R., and Kretsinger, R. (2003) The high-resolution structure of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase reveals a twist in the plane of bound phosphoenolpyruvate. *Biochemistry 42*, 3766–3776.

[28] Schoner, R., and Herrmann, K. (1976) 3-Deoxy-D-*arabino*-heptulosonate

7-phosphate synthase. Purification, properties, and kinetics of the tyrosine-sensitive isoenzyme from *Escherichia coli*. *J. Biol. Chem. 251*, 5440–5447.

[29] Simpson, R., and Davidson, B. (1976) Studies on 3-deoxy-D-*arabino*heptulosonate-7-phosphate synthetase (phe) from *Escherichia-coli*-K12. Kinetic properties. *Eur. J. Biochem. 70*, 501–507.

[30] Staub, M., and Denes, G. (1969) Purification and properties of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate-synthase (phenylalanine sensitive) of *Escherichia coli* K12 Inhibition of activity of enzyme with phenylalanine and functional group-specific reagents. *Biochim. Biophys. Acta 178*, 599–608.

[31] McCandliss, R., Poling, M., and Herrmann, K. (1978) 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate synthase – Purification and molecular characterization of phenylalanine-sensitive isoezyme from *Escherichia-coli*. *J. Biol. Chem. 253*, 4259–4265.

[32] Hu, C., Jiang, P., Xu, J., Wu, Y., and Huang, W. (2003) Mutation analysis of the feedback inhibition site of phenylalanine-sensitive 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase of *Escherichia coli*. *J. Basic Microb. 43*, 399–406.

[33] Hartmann, M., Schneider, T., Pfeil, A., Heinrich, G., Lipscomb, W., and Braus, G. (2003) Evolution of feedback-inhibited/barrel isoenzymes by gene duplication and a single mutation. *Proc. Natl. Acad. Sci. U. S. A. 100*, 862–867.

[34] Kónig, V., Pfeil, A., Braus, G., and Schneider, T. (2004) Substrate and metal complexes of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Saccharomyces cerevisiae* provide new insights into the catalytic mechanism. *J. Mol. Biol. 337*, 675–690.

[35] Subramaniam, P., Xie, G., Xia, T., and Jensen, R. (1998) Substrate ambiguity of 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase from *Neisseria gonorrhoeae* in the context of its membership in a protein fam-

ily containing a subset of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthases. *J. Bacteriol. 180*, 119–127.

[36] Raetz, C. (1990) Biochemistry of endotoxins. *Annu. Rev. Biochem. 59*, 129–170.

[37] Rick, P., and Osborn, M. (1977) Lipid-A mutants of *Salmonella-typhimurium*-Characterization of a conditional lethal mutant in 3-deoxy-D-*manno*-octulosonate-8-phosphate synthetase. *J. Biol. Chem. 252*, 4895–4903.

[38] Schofield, L., Patchett, M., and Parker, E. (2004) Expression, purification, and characterization of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Pyrococcus furiosus*. *Protein Expression Purif. 34*, 17–27.

[39] Shumilin, I., Bauerle, R., Wu, J., Woodard, R., and Kretsinger, R. (2004) Crystal structure of the reaction complex of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Thermotoga maritima* refines the catalytic mechanism and indicates a new mechanism of allosteric regulation. *J. Mol. Biol. 341*, 455–466.

[40] Zhou, L., Wu, J., Janakiraman, V., Shumilin, I. A., Bauerle, R., Kretsinger, R. H., and Woodard, R. W. (2011) Structure and characterization of the 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Aeropyrum pernix*. *Bioorg. Chem. 40*, 79–86.

[41] Fowden, L. (1965) The chemical approach to plants. *Sci. Progr. 53*, 583–599.

[42] Gosset, G., Bonner, C., and Jensen, R. (2001) Microbial origin of plant-type 2-keto-3-deoxy-D-*arabino*–heptulosonate 7-phosphate synthases, exemplified by the chorismate-and tryptophan-regulated enzyme from *Xanthomonas campestris*. *J. Bacteriol. 183*, 4061–4070.

[43] Kloosterman, H., Hessels, G., Vrijbloed, J., Euverink, G., and Dijkhuizen, L. (2003) (De)regulation of key enzyme steps in the

shikimate pathway and phenylalanine-specific pathway of the actinomycete *Amycolatopsis methanolica*. *Microbiol. - SGM 149*, 3321–3330.

[44] Silakowski, B., Kunze, B., and Muller, R. (2000) *Stigmatella aurantiaca* Sg a15 carries genes encoding type I and type II 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthases: involvement of a type II synthase in aurachin biosynthesis. *Arch. Microbiol. 173*, 403–411.

[45] Guo, J., and Frost, J. (2002) Kanosamine biosynthesis: A likely source of the aminoshikimate pathway's nitrogen atom. *J. Am. Chem. Soc. 124*, 10642–10643.

[46] Webby, C., Baker, H., Lott, J., Baker, E., and Parker, E. (2005) The structure of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Mycobacterium tuberculosis* reveals a common catalytic scaffold and ancestry for type I and type II enzymes. *J. Mol. Biol. 354*, 927–939.

[47] Webby, C., Patchett, M., and Parker, E. (2005) Characterization of a recombinant type II 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Helicobacter pylori*. *Biochem. J. 390*, 223–230.

[48] Schofield, L., Anderson, B., Patchett, M., Norris, G., Jameson, G., and Parker, E. (2005) Substrate ambiguity and crystal structure of *Pyrococcus furiosus* 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase: An ancestral 3-deoxyald-2-ulosonate-phosphate synthase? *Biochemistry 44*, 11950–11962.

[49] Williamson, R., Pietersma, A., Jameson, G., and Parker, E. (2005) Stereospecific deuteration of 2-deoxyerythrose 4-phosphate using 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. *Bioorg. Med. Chem. Lett. 15*, 2339–2342.

[50] DeLeo, A., Dayan, J., and Sprinson, D. (1973) Purification and kinetics of tyrosine-sensitive 3-deoxy-D-*arabino*-heptulosonic acid 7-phosphate synthetase from *Salmonella*. *J. Biol. Chem. 248*, 2344–2353.

[51] DeLeo, A., and Sprinson, D. (1968) Mechanism of 3-deoxy-D-*arabino*-

heptulosonate 7-phosphate (DAHP) synthetase. *Biochem. Biophys. Res. Commun. 32*, 873–877.

[52] Dotson, G., Dua, R., Clemens, J., Wooten, E., and Woodard, R. (1995) Overproduction and one-step purification of *Escherichia-coli* 3-deoxy-D-*manno*-octulosonic acid 8-phosphate synthase and oxygen-transfer studies during cataylsis using isotopic-shifted heteronuclear NMR. *J. Biol. Chem. 270*, 13698–13705.

[53] Staub, M., and Denes, G. (1967) A kinetic study of the mechanism of action of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase in *Escherichia coli* K12. *Biochim. Biophys. Acta 132*, 528–530.

[54] Lambert, J., Boocock, M., and Coggins, J. (1985) The 3-dehydroquinate synthase activity of the pentafunctional arom enzyme complex of *Neurospora-crassa* is Zn^{2+}-dependent. *Biochem. J. 226*, 817–829.

[55] Paravicini, G., Schmidheini, T., and Braus, G. (1989) Purification and properties of the 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase (phenylalanine-inhibitable) of *Saccharomyces-cervisiae. Eur. J. Biochem. 186*, 361–366.

[56] Ray, J., and Bauerle, R. (1991) Purification and properties of tryptophan-sensitive 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia-coli. J. Bacteriol. 173*, 1894–1901.

[57] Ganson, R., and Jensen, R. (1988) The essential role of cobalt in the inhibition of the cytosolic isozyme of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Nicotiana-silvestris* by glyphosate. *Arch. Biochem. Biophys. 260*, 85–93.

[58] Baasov, T., and Knowles, J. (1989) Is the 1st enzyme of the shikimate pathway, 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase (tyrosine sensitive), a copper metalloenzyme? *J. Bacteriol. 171*, 6155–6160.

[59] Schnappauf, G., Hartmann, M., Kunzler, M., and Braus, G. (1998) The two 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase isoenzymes

from *Saccharomyces cerevisiae* show different kinetic modes of inhibition. *Arch. Microbiol. 169*, 517–524.

[60] Stephens, C., and Bauerle, R. (1991) Analysis of the metal requirement of 3-deoxy-D-*arabino*-hepulosonate-7-phosphate synthase from *Escherichia-coli*. *J. Biol. Chem. 266*, 20810–20817.

[61] Staub, M., and Denes, G. (1969) Purification and properties of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase (phenylalanine sensitive) of *Escherichia coli* K12. Purification of enzyme and some of its catalytic properties. *Biochim. Biophys. Acta 178*, 588–598.

[62] McCandliss, R., and Herrmann, K. (1978) Iron, an essential element for biosynthesis of aromatic-compounds. *Proc. Natl. Acad. Sci. U. S. A. 75*, 4810–4813.

[63] Li, Y., and Evans, J. (1996) The hard-soft acid-base principle in enzymatic catalysis: Dual reactivity of phosphoenolpyruvate. *Proc. Natl. Acad. Sci. U. S. A. 93*, 4612–4616.

[64] Wu, J., Howe, D., and Woodard, R. (2003) *Thermotoga maritima* 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAHP) synthase. *J. Biol. Chem. 278*, 27525–27531.

[65] Wierenga, R. (2001) The TIM-barrel fold: a versatile framework for efficient enzymes. *FEBS Lett. 492*, 193–198.

[66] Webby, C., Jiao, W., Hutton, R., Blackmore, N., Baker, H., Baker, E., Jameson, G., and Parker, E. (2010) Synergistic allostery, a sophisticated regulatory network for the control of aromatic amino acid biosynthesis in *Mycobacterium tuberculosis*. *J. Biol. Chem. 285*, 30567–30576.

[67] LeMarechal, P., Froussios, C., Level, M., and Azerad, R. (1980) Enzymic properties of phosphanic analogs of D-erythrose 4-phosphate. *Biochem. Biophys. Res. Commun. 92*, 1097–1103.

[68] Sheflyan, G., Howe, D., Wilson, T., and Woodard, R. (1998) Enzymatic synthesis of 3-deoxy-D-*manno*-octulosonate 8-phosphate, 3-

deoxy-D-*altro*-octulosonate 8-phosphate, 3,5-dideoxy-D-*gluco*(*manno*)-octulosonate 8-phosphate by 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. *J. Am. Chem. Soc. 120*, 11027–11032.

[69] Tran, D., Pietersma, A. L., Schofield, L. R., Rost, M., Jameson, G. B., and Parker, E. J. (2011) Investigating the role of the hydroxyl groups of substrate erythrose 4-phosphate in the reaction catalysed by the first enzyme of the shikimate pathway. *Bioorg. Med. Chem. Lett. 21*, 6838–6841.

[70] Ahn, M., Pietersma, A., Schofield, L., and Parker, E. (2005) Mechanistic divergence of two closely related aldol-like enzyme-catalysed reactions. *Org. Biomol. Chem. 3*, 4046–4049.

[71] Bell, G., Burant, C., Takeda, J., and Gould, G. (1993) Structure and function of mammalian facilitative sugar transporters. *J. Biol. Chem. 268*, 19161–19164.

[72] Bouche, C., Serdy, S., Kahn, C., and Goldfine, A. (2004) The cellular fate of glucose and its relevance in type 2 diabetes. *Endocr. Rev. 25*, 807–830.

[73] Sakami, W., and Harrington, H. (1963) Amino acid metabolism. *Annu. Rev. Biochem. 32*, 355–398.

[74] Braus, G. (1991) Aromatic amino-acid biosynthesis in the yeast *Saccharomyces-cerevisiae*: A model system for the regulation of a eukaryotic biosynthetic pathway. *Microbiol. Rev. 55*, 349–370.

[75] Faergeman, N., and Knudsen, J. (1997) Role of long-chain fatty acyl-CoA esters in the regulation of metabolism and in cell signalling. *Biochem. J. 323*, 1–12.

[76] Doble, B., and Woodgett, J. (2003) GSK-3: Tricks of the trade for a multi-tasking kinase. *J. Cell Sci. 116*, 1175–1186.

[77] Carr, C., and Kim, P. (1993) A spring-loaded mechanism for the con-

formational change of influenza hemagglutinin. *Cell (Cambridge, MA, United States) 73*, 823–832.

[78] Helmstaedt, K., Krappmann, S., and Braus, G. (2001) Allosteric regulation of catalytic activity: *Escherichia coli* aspartate transcarbamoylase versus yeast chorismate mutase. *Microbiol. Mol. Biol. Rev. 65*, 404–421.

[79] Im, S., Davidson, H., and Pittard, J. (1971) Phenylalanine and tyrosine biosynthesis in *Escherichia-coli* K-12 - mutants derepressed for 3-deoxy-D-*arabino*heptulosonic acid 7-phosphate synthetase (phe), 3-deoxy-D-*arabino*heptulosonic acid 7-phosphate synthetase (tyr), chorismate mutase T-prephenate dehydrogenase, and transaminase-A. *J. Bacteriol. 108*, 400–409.

[80] Brown, K., and Somervil, R. (1971) Repression of aromatic amino acid biosynthesis in *Escherichia-coli* K-12. *J. Bacteriol. 108*, 386–399.

[81] Tribe, D., Camakaris, H., and Pittard, J. (1976) Constitutive and repressible enzymes of the common pathway of aromatic biosynthesis in *Escherichia coli* K-12: regulation of enzyme synthesis at different growth rates. *J. Bacteriol. 127*, 1085–1097.

[82] Grove, C., and Gunsalus, R. (1987) Regulation of the aroH operon of *Escherichia-coli* by the tryptophan repressor. *J. Bacteriol. 169*, 2158–2164.

[83] Zurawski, G., Gunsalus, R., Brown, K., and Yanofsky, C. (1981) Structure and regulation of *aro*H, the structural gene for the tryptophan-repressible 3-deoxy-D-*arabino*-heptulosonic acid 7-phosphate synthetase of *Escherichia-coli*. *J. Mol. Biol. 145*, 47–73.

[84] Arvidson, D., Bruce, C., and Gunsalus, R. (1986) Interaction of the *Escherichia-coli* trp aporepressor with its ligand, L-tryptophan. *J. Biol. Chem. 261*, 238–243.

[85] Monod, J., Wyman, J., and Changeux, J. (1965) On nature of allosteric transitions - A plausible model. *J. Mol. Biol. 12*, 88–118.

[86] Bohr, C. (1904) Die Saurestoffe des genuinen Blutfarbstoffes und des aus dem Blute darstellten Hämoglobins. *Centralbl. für Physiol. 23*.

[87] Cori, G., Colowick, S., and Cori, C. (1938) The action of nucleotides in the disruptive phosphorylation of glycogen. *J. Biol. Chem. 123*, 381–389.

[88] Hill, A. (1910) The possible effect of the aggregation of the molecules of haemoglobin on its dissociation curve. *J. Physiol. 40*, 4–7.

[89] Novick, A., and Szilard, L. In *Dynamics of growth processes*; Boell, E., Ed.; Princeton University Press, Princeton, New Jersey, 1954; pp 21–32, Experiments with the chemostat on the rates of amino acid synthesis in bacteria.

[90] Umbarger, H. (1961) Feedback control by endproduct inhibition. *Cold Spr. Harb. Symp. Quant. Biol. 26*, 301–312.

[91] Jacob, J., F; Monod (1961) On regulation of gene activity. *Cold Spr. Harb. Symp. Quant. Biol. 26*, 193–209.

[92] Changeux, J. (1961) Feedback control mechanism of biosynthetic L-threonine deaminase by L-isoleucine. *Cold Spr. Harb. Symp. Quant. Biol. 26*, 313–318.

[93] Changeux, J. (1963) Allosteric interactions on biosynthetic L-threonine deaminase from *E.coli* K12. *Cold Spr. Harb. Symp. Quant. Biol. 28*, 497–504.

[94] Koshland, D., Nemethy, G., and Filmer, D. (1966) Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry 5*, 365–385.

[95] Garrett., R., and Grisham, C. *Biochemistry*, third edition ed.; Belmont: Thomson Brooks/Cole, 2005.

[96] Mayer, M., Schroder, H., Rudiger, S., Paal, K., Laufen, T., and Bukau, B. (2000) Multistep mechanism of substrate binding determines chaperone activity of Hsp70. *Nat. Struct. Biol. 7*, 586–593.

[97] Johnson, L. (1994) Control by protein-phosphorylation. *Nat. Struct. Biol. 1*, 657–659.

[98] Volkman, B., Lipson, D., Wemmer, D., and Kern, D. (2001) Two-state allosteric behavior in a single-domain signaling protein. *Science 291*, 2429–2433.

[99] Swain, J., and Gierasch, L. (2006) The changing landscape of protein allostery. *Curr. Opin. Struct. Biol. 16*, 102–108.

[100] Suel, G., Lockless, S., Wall, M., and Ranganathan, R. (2003) Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nat. Struct. Biol. 10*, 59–69.

[101] Tsai, C.-J., del Sol, A., and Nussinov, R. (2008) Allostery: Absence of a change in shape does not imply that allostery is not at play. *J. Mol. Biol. 378*, 1–11.

[102] Cooper, A., and Dryden, D. (1984) Allostery without conformational change - A plausible model. *Eur. Biophys. J. Biophy. 11*, 103–109.

[103] Walker, G., Dunbar, B., Hunter, S., Nimmo, H., and Coggins, J. (1996) Evidence for a novel class of microbial 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase in *Streptomyces coelicolor* A3 (2), *Streptomyces rimosus* and *Neurospora crassa*. *Microbiology 142*, 1973–1982.

[104] Shumilin, I., Zhao, C., Bauerle, R., and Kretsinger, R. (2002) Allosteric inhibition of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase alters the coordination of both substrates. *J. Mol. Biol. 320*, 1147–1156.

[105] Wu, J., and Woodard, R. (2006) New insights into the evolutionary links relating to the 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase subfamilies. *J. Biol. Chem. 281*, 4042–4048.

[106] Jiao, W., Hutton, R. D., Cross, P. J., Jameson, G. B., and Parker, E. J. (2011) Dynamic cross-talk among remote binding sites: The molecular basis for unusual synergistic allostery. *J. Mol. Biol. 415*, 716–726.

[107] Umbarger, H. (1978) Amino-acid biosynthesis and its regulation. *Annu. Rev. Biochem. 47*, 533–606.

[108] Sasso, S., Ökvist, M., Roderer, K., Gamper, M., Codoni, G., Krengel, U., and Kast, P. (2009) Structure and function of a complex between chorismate mutase and DAHP synthase: efficiency boost for the junior partner. *EMBO J. 28*, 2128–2142.

[109] Euverink, G., Hessels, G., Franke, C., and Dijkhuizen, L. (1995) Chorismate mutase and 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase of the methylotrophic actinomycete *Amycolatopsis-methanolica*. *Appl. Environ. Microbiol. 61*, 3796–3803.

[110] Yoo, J., Bang, H., Lee, E., Kim, S., and Lee, J. (1993) Purification and characterisation of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate-synthetase from *Streptomyces caespitosus*. *Korean J. Microbiol. 6*, 340–345.

[111] Stetter, K. O. (1996) Hyperthermophilic procaryotes. *FEMS Microbiol. Rev. 18*, 149–158.

[112] Woese, C., Ksndler, O., and Wheelis, M. (1990) Towards a natural system of organisms - Proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. U. S. A. 87*, 4576–4579.

[113] Huber, R., Langworthy, T., Konig, H., Thomm, M., Woese, C., Sleytr, U., and Stetter, K. (1986) *Thermotoga-maritima* Sp-Nov represents a new genus of extremely thermophillic Eubacteria growing up to 90 °C. *Arch. Microbiol. 144*, 324–333.

[114] Achenbachrichter, L., Gupta, R., Stetter, K., and Woese, C. (1987) Were the original Eubacteria thermophiles? *Syst. Appl. Microbiol. 9*, 34–39.

[115] Bocchetta, M., Gribaldo, S., Sanangelantoni, A., and Cammarano, P. (2000) Phylogenetic depth of the bacterial genera Aquifex and Thermotoga inferred from analysis of ribosomal protein, elongation factor, and RNA polymerase subunit sequences. *J. Mol. Evol. 50*, 366–380.

[116] Tabor, S., and Richardson, C. (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. U. S. A. 82*, 1074–1078.

[117] Costantino, H., Brown, S., and Kelly, R. (1990) Purification and characterization of an alpha-glucosidase fropm a hyperthermophilic Archebacterium, *Pyrococcus-furiosus*, exhibiting a temperature of optimum 105–115 degrees C. *J. Bacteriol. 172*, 3654–3660.

[118] Percival, R., Devine, D., Bonass, W., Marsh, P., and Curtis, M. (1999) The effect of temperature on LPS expression by *Porphyromonas gingivalis*. *J. Dent. Res. 78*, 1070.

[119] David H. Bergey., D. R. B. *Bergey's manual of systematic bacteriology*, second edition ed.; Springer, 2009; Vol. Volume Three.

[120] Sako, Y., Nomura, N., Uchida, A., Ishida, Y., Morii, H., Koga, Y., Hoaki, T., and Maruyama, T. (1996) *Aeropyrum pernix* gen. nov., sp. nov., a novel aerobic hyperthermophilic archaeon growing at temperatures up to 100 degrees C. *Int. J. Syst. Bacteriol. 46*, 1070–1077.

[121] Gorisch, H., and Lingens, F. (1971) 3-Deoxy-D-*arabino*-heptulosonate-7-phosphate synthase of the *Streptomyces-aureofaciens* tü 24 .I. Partial purification and properties. *Biochim. Biophys. Acta 242*, 630–636.

[122] Jossek, R., Bongaerts, J., and Sprenger, G. (2001) Characterization of a new feedback-resistant 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase AroF of *Escherichia coli*. *FEMS Microbiol. Lett. 202*, 145–148.

[123] Patchett, M., Neal, T., Schofield, L., Strange, R., Daniel, R., and Morgan, H. (1989) Heat treatment purification of thermostable cellulase and hemicellulase enzymes expressed in *E. coli*. *Enzyme Microb. Technol. 11*, 113 –115.

[124] Waugh, D. (2005) Making the most of affinity tags. *Trends Biotechnol. 23*, 316–320.

[125] Svergun, D. (1992) Determination of the regularization parameter in

indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr. 25*, 495–503.

[126] Menegay, H., Myers, M., Moeslein, F., and Landreth, G. (2000) Biochemical characterization and localization of the dual specificity kinase CLK1. *J. Cell Sci. 113*, 3241–3253.

[127] Akama, K., and Takaiwa, F. (2007) C-terminal extension of rice glutamate decarboxylase (OsGAD2) functions as an autoinhibitory domain and overexpression of a truncated mutant results in the accumulation of extremely high levels of GABA in plant cells. *J. Exp. Bot. 58*, 2699–2707.

[128] Bell, J. K., Pease, P. J., Bell, J. E., Grant, G. A., and Banaszak, L. J. (2002) De-regulation of D-3-phosphoglycerate dehydrogenase by domain removal. *Eur. J. Biochem. 269*, 4176–4184.

[129] Cross, P. J., Dobson, R. C. J., Patchett, M. L., and Parker, E. J. (2011) Tyrosine latching of a regulatory gate affords allosteric control of aromatic amino acid biosynthesis. *J. Biol. Chem. 286*, 10216–10224.

[130] Grant, G. (2006) The ACT domain: a small molecule binding domain and its role as a common regulatory element. *J. Biol. Chem. 281*, 33825–33829.

[131] Curien, G., Biou, V., Mas-Droux, C., Robert-Genthon, M., Ferrer, J., and Dumas, R. (2008) Amino acid biosynthesis: new architectures in allosteric enzymes. *Plant Physiol. Biochem. 46*, 325–339.

[132] Aravind, L., and Koonin, E. (1999) Gleaning non-trivial structural, functional and evolutionary information about proteins by iterative database searches. *J. Mol. Biol. 287*, 1023–1040.

[133] Schuller, D., Grant, G., and Banaszak, L. (1995) The allosteric ligand site in the V-max-type cooperative enzyme phosphoglycerate dehydrogenase. *Nat. Struct. Biol. 2*, 69–76.

[134] Steller, I., Bolotovsky, R., and Rossmann, M. (1997) An algorithm for

automatic indexing of oscillation images using Fourier analysis. *J. Appl. Crystallogr. 30*, 1036–1040.

[135] Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr, Sect D: Biol Crystallogr 62*, 72–82.

[136] Laskowski, R., Gerick, F., and Thornton, J. (2009) The structural basis of allosteric regulation in proteins. *FEBS Lett. 583*, 1692–1698.

[137] Liberles, J., Thorolfsson, M., and Martinez, A. (2005) Allosteric mechanisms in ACT domain containing enzymes involved in amino acid metabolism. *Amino Acids 28*, 1–12.

[138] Cho, Y., Sharma, V., and Sacchettini, J. (2003) Crystal structure of ATP phosphoribosyltransferase from *Mycobacterium tuberculosis*. *J. Biol. Chem. 278*, 8333–8339.

[139] Zhang, S., Pohnert, G., Kongsaeree, P., Wilson, D., Clardy, J., and Ganem, B. (1998) Chorismate mutase-prephenate dehydratase from *Escherichia coli*. *J. Biol. Chem. 273*, 6248–6253.

[140] Svergun, D., and Koch, M. (2003) Small-angle scattering studies of biological macromolecules in solution. *Rep. Prog. Phys. 66*, 1735–1782.

[141] Liao, H., Lin, L., Chien, H., and Hsu, W. (2001) Serine 187 is a crucial residue for allosteric regulation of *Corynebacterium glutamicum* 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase. *FEMS Microbiol. Lett. 194*, 59–64.

[142] Kossiakoff, A., Shpungin, J., and Sintchak, M. (1990) Hydroxyl hydrogen conformations in trypsin determined by neutron-diffraction solvent difference map method - Relative importance of steric and electrostatic factors in defining hydrogen-bonding geometries. *Proc. Natl. Acad. Sci. U. S. A. 87*, 4468–4472.

[143] Blessing, R., and McGandy, E. (1972) Base stacking and hydrogen-bonding in crystals of imidazolium dihydrogen orthophosphate. *J. Am. Chem. Soc. 94*, 4034–4035.

[144] McGaughey, G., Gagne, M., and Rappe, A. (1998) π-stacking interactions - Alive and well in proteins. *J. Biol. Chem. 273*, 15458–15463.

[145] Abraham, R., Eivazi, F., Pearson, H., and Smith, K. (1976) π-π Aggregation in metalloporphyrins - Causative factors. *J. Chem. Soc. Chem. Comm.* 699–701.

[146] Shetty, A., Zhang, J., and Moore, J. (1996) Aromatic pi-stacking in solution as revealed through the aggregation of phenylacetylene macrocycles. *J. Am. Chem. Soc. 118*, 1019–1027.

[147] Tanner, D., Fitzgerald, J., and Phillips, B. (1989) The kevlar story - An advanced materials case-study. *Angew. Chem. Int. Edit. 28*, 649–654.

[148] Ishida, T., Doi, M., Ueda, H., Inoue, M., and Scheldrick, G. (1988) Structural studies of the interaction between indole-derivatives and biologically important aromatic-compounds. 20. Specific ring stacking interaction on the tryptophan 7-methylguanine system – Comparative crystallographic studies of indole derivatives 7-methylguanine base, nucleoside, and nucleotide complexes. *J. Am. Chem. Soc. 110*, 2286–2294.

[149] Kamiichi, K., Danshita, M., Minamino, N., Doi, M., Ishida, T., and Inoue, M. (1986) Indole ring binds to 7-methylguanine base by pi-pi stacking interaction - Crystal structure of 7-methylguanosine 5'-monophosphate-tryptamine complex. *FEBS Lett. 195*, 57–60.

[150] Burley, S., and Petsko, G. (1985) Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science 229*, 23–28.

[151] Jasanoff, A., and Weiss, M. (1993) Aromatic histidine interactions in the zinc finger motif - Structural in equivalence of phenylalanine and tyrosine in the hydrophobic core. *Biochemistry 32*, 1423–1432.

[152] Ostermeier, M., and Benkovic, S. *Adavnces in Protein Chemistry*; 2001; Vol. 55; pp 29–77.

[153] Crameri, A., Raillard, S., Bermudez, E., and Stemmer, W. (1998) DNA

shuffling of a family of genes from diverse species accelerates directed evolution. *Nature 391*, 288–291.

[154] Stemmer, W. (1994) DNA shuffling by random fragmentation and reassembly — In-vitro recombination for molecular evolution. *Proc. Natl. Acad. Sci. U. S. A. 91*, 10747–10751.

[155] Nixon, A., Ostermeier, M., and Benkovic, S. (1998) Hybrid enzymes: manipulating enzyme design. *Trends Biotechnol. 16*, 258–264.

[156] Ettema, T., Brinkman, A., Tani, T., Rafferty, J., and van der Oost, J. (2002) A novel ligand-binding domain involved in regulation of amino acid metabolism in prokaryotes. *J. Biol. Chem. 277*, 37464–37468.

[157] Chipman, D., and Shaanan, B. (2001) The ACT domain family. *Curr. Opin. Struct. Biol. 11*, 694–700.

[158] Kobe, B., Jennings, I., House, C., Michell, B., Goodwill, K., Santarsiero, B., Stevens, R., Cotton, R., and Kemp, B. (1999) Structural basis of autoregulation of phenylalanine hydroxylase. *Nat. Struct. Biol. 6*, 442–448.

[159] de Marco, A. (2007) Protocol for preparing proteins with improved solubility by co-expressing with molecular chaperones in *Escherichia coli*. *Nat. Protoc. 2*, 2632–2639.

[160] Allison, T. M., Yeoman, J. A., Hutton, R. D., Cochrane, F. C., Jameson, G. B., and Parker, E. J. (2010) Specificity and mutational analysis of the metal-dependent 3-deoxy-D-*manno*-octulosonate 8-phosphate synthase from *Acidithiobacillus ferrooxidans*. *Biochim. Biophys. Acta, Proteins Proteomics 1804*, 1526–1536.

[161] Hedstrom, L., Szilagyi, L., and Rutter, W. (1992) Converting trypsin to chymotrypsin - The role of surface loops. *Science 255*, 1249–1253.

[162] Venekei, I., Szilagyi, L., Graf, L., and Rutter, W. (1996) Attempts to convert chymotrypsin to trypsin. *FEBS Lett. 379*, 143–147.

[163] Wilks, H., Hart, K., Feeney, R., Dunn, C., Muirhead, H., Chia, W.,

Barstow, D., Atkinson, T., Clarke, A., and Holbrook, J. (1988) A specific, highly-active malate-dehydrogenase by redesign of a lactate-dehydrogenase framework. *Science 242*, 1541–1544.

[164] Wells, J., Cunningham, B., Graycar, T., and Estell, D. (1987) Recruitment of the substrate-specificity properties from one enzyme into a related one by protein engineering. *Proc. Natl. Acad. Sci. U. S. A. 84*, 5167–5171.

[165] Ballinger, M., Tom, J., and Wells, J. (1995) Designing subtilisin BPN' to cleave substrates containing dibasic residues. *Biochemistry 34*, 13312–13319.

[166] Ballinger, M., Tom, J., and Wells, J. (1996) Furilisin: A variant of subtilisin BPN' engineered for cleaving tribasic substrates. *Biochemistry 35*, 13579–13585.

[167] Wilks, H., Halsall, D., Atkinson, T., Chia, W., Clarke, A., and Holbrook, J. (1990) Designs for a broad substrate-specificity keto acid dehydrogenase. *Biochemistry 29*, 8587–8591.

[168] Wolfson, A., Kanaoka, M., Lau, F., and Ringe, D. (1991) Insertion of an elastase-binding loop into interleukin-1-beta. *Protein Eng. 4*, 313–317.

[169] Brennan, C., Christianson, K., Surowy, T., and Mandecki, W. (1994) Modulation of enzyme-activity by antibody-binding to an alkaline-phosphatase-epitope hybrid protein. *Protein Eng. 7*, 509–514.

[170] Singh, A., and Hayashi, K. (1995) Construction of beta-glucosidases with improved enzymatic–properties. *J. Biol. Chem. 270*, 21928–21933.

[171] Doi, N., and Yanagawa, H. (1999) Design of generic biosensors based on green fluorescent proteins with allosteric sites by directed evolution. *FEBS Lett. 453*, 305–307.

[172] Guntas, G., and Ostermeier, M. (2004) Creation of an allosteric enzyme by domain insertion. *J. Mol. Biol. 336*, 263–273.

[173] Baird, G., Zacharias, D., and Tsien, R. (1999) Circular permutation and

receptor insertion within green fluorescent proteins. *Proc. Natl. Acad. Sci. U. S. A. 96*, 11241–11246.

[174] Siegel, M., and Isacoff, E. (1997) A genetically encoded optical probe of membrane voltage. *Neuron 19*, 735–741.

[175] Fehr, M., Frommer, W., and Lalonde, S. (2002) Visualization of maltose uptake in living yeast cells by fluorescent nanosensors. *Proc. Natl. Acad. Sci. U. S. A. 99*, 9846–9851.

[176] De Lorimier, R., Smith, J., Dwyer, M., Looger, L., Sali, K., Paavola, C., Rizk, S., Sadigov, S., Conrad, D., Loew, L., and Hellinga, H. (2002) Construction of a fluorescent biosensor family. *Protein Sci. 11*, 2655–2675.

[177] Stayton, P., Shimoboji, T., Long, C., Chilkoti, A., Chen, G., Harris, J., and Hoffman, A. (1995) Control of protein-ligand recognition using a stimuli-responsive polymer. *Nature 378*, 472–474.

[178] Rivera, V. (1998) Controlling gene expression using synthetic ligands. *Methods 14*, 421–429.

[179] Picard, D. *Applications of chimeric genes and hybrid proteins pt B*; Method. Enzymol.; 2000; Vol. 327; pp 385–401.

[180] Guo, Z., Zhou, D., and Schultz, P. (2000) Designing small-molecule switches for protein-protein interactions. *Science 288*, 2042–2045.

[181] Pelletier, J., Campbell-Valois, F., and Michnick, S. (1998) Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments. *Proc. Natl. Acad. Sci. U. S. A. 95*, 12141–12146.

[182] Betton, J., Jacob, J., Hofnung, M., and Broome-Smith, J. (1997) Creating a bifunctional protein by insertion of beta-lactamase into the maltodextrin-binding protein. *Nat. Biotechnol. 15*, 1276–1279.

[183] Fushinobu, S., Ohta, T., and Matsuzawa, H. (1998) Homotropic activation via the subunit interaction and allosteric symmetry revealed on

analysis of hybrid enzymes of L-lactate dehydrogenase. *J. Biol. Chem. 273*, 2971–2976.

[184] Remy, I., and Michnick, S. (1999) Clonal selection and in vivo quantitation of protein interactions with protein-fragment complementation assays. *Proc. Natl. Acad. Sci. U. S. A. 96*, 5394–5399.

[185] Dueber, J., Yeh, B., Chak, K., and Lim, W. (2003) Reprogramming control of an allosteric signaling switch through modular recombination. *Science 301*, 1904–1908.

[186] Sharff, A., Rodseth, L., Spurlino, J., and Quiocho, F. (1992) Crystallographic evidence of a large-ligand induced hinge-twist motion between the two domains of the maltodextrin binding-protein involved in active-transport and chemotaxis. *Biochemistry 31*, 10657–10663.

[187] Schrödinger, LLC, The PyMOL Molecular Graphics System, Version 1.5. 2010.

[188] Laemmli, U. (1970) Cleavage of structural proteins during assembly of head of bacteriophage-T4. *Nature 227*, 680–685.

[189] Bradford, M. (1976) Rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding. *Anal. Biochem. 72*, 248–254.

[190] Storoni, L., McCoy, A., and Read, R. (2004) Likelihood-enhanced fast rotation functions. *Acta Crystallogr , Sect D: Biol Crystallogr 60*, 432–438.

[191] Bailey, S. (1994) The CCP4 Suite - Programs for protein crystallography. *Acta Crystallogr , Sect D: Biol Crystallogr 50*, 760–763.

[192] Adams, P., Afonine, P., BunkÃşczi, G., Chen, V., Davis, I., Echols, N., Headd, J., Hung, L., Kapral, G., and Grosse-Kunstleve, R. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 66*, 213–221.

[193] Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 2126–2132.

[194] Painter, J., and Merritt, E. (2006) TLSMD web server for the generation of multi-group TLS models. *J. Appl. Crystallogr. 39*, 109–111.

[195] Chen, V., Arendall, W., 3rd Headd, J., Keedy, D., Immormino, R., Kapral, G., Murray, L., Richardson, J., and Richardson, D. (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 66*, 12–21.

[196] Svergun, D., Barberato, C., and Koch, M. (1995) CRYSOL - Program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr. 28*, 768–773.

[197] Mertens, H., and Svergun, D. (2010) Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *J. Struct. Biol. 172*, 128–141.

[198] Mills, R., Trewhella, J., Qiu, T., Welte, T., Ryan, T., Hanley, T., Knott, R., Lithgow, T., and Mulhern, T. (2009) Domain organization of the monomeric form of the Tom70 mitochondrial import receptor. *J. Mol. Biol. 388*, 1043–1058.